NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

THE USE OF CONDITIONAL AND POTENTIAL INSTABILITY AXES FOR SEVERE WEATHER FORECASTING

by

Douglas A. Oltmer

March 2009

Thesis Advisor: Wendell Nuss
Second Reader: Anthony Eckel

Approved for public release; distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2009	Master's Thesis

4. TITLE AND SUBTITLE The Use of Conditional and Potential Instability Axes for Severe Weather Forecasting	5. FUNDING NUMBERS
6. AUTHOR(S) Douglas A. Oltmer	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution is unlimited.	

13. ABSTRACT (maximum 200 words)

In this thesis, we used an innovative approach to severe weather forecasting using traditional forecast parameters. The maximum theta-e (potential instability) in the lowest 300 mb and the deep-layer lapse rate (conditional instability) fields are two commonly used forecast parameters for severe weather forecasting, and the ridges of these fields, or more specifically, the ridge intersections, were hypothesized to indicate regions for enhanced severe convection. In addition, the sharpness of these ridge axes may correlate to an increase in severe potential. To test this theory, a mathematical formula was devised to quantitatively assess ridge strength. Then, using theta-e, lapse rate and their respective ridge strengths as predictors, a linear discriminate analysis was performed on dependent and independent datasets from the spring of 2008. Severe probabilistic forecasts were produced using the discriminate analysis and verified using two independent methods. Skill metrics calculated for the forecasts determined there is significant positive skill of the forecast technique. As a side note, this research attempted to determine the wavelength features (meso-α to meso-β) that recorded the highest level of skill using this method. No significant difference was noted in the different wavelength feature forecasts, possibly due to the verification method.

14. SUBJECT TERMS Severe Weather, Conditional Instability, Potential Instability, Discriminate Analysis, Theta-e, Lapse Rate, Ridge Axis Formula	15. NUMBER OF PAGES 115
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited.

THE USE OF CONDITIONAL AND POTENTIAL INSTABILITY AXES FOR SEVERE WEATHER FORECASTING

Douglas A. Oltmer
Captain, United States Air Force
B.S., Texas A&M University, 2000
B.S., Texas A&M University, 2003

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
March 2009

Author: Douglas A. Oltmer

Approved by: Wendell Nuss
Thesis Advisor

Anthony Eckel, Maj., USAF
Second Reader

Philip A. Durkee
Chairman, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

In this thesis, we used an innovative approach to severe weather forecasting using traditional forecast parameters. The maximum theta-e (potential instability) in the lowest 300 mb and the deep-layer lapse rate (conditional instability) fields are two commonly used forecast parameters for severe weather forecasting, and the ridges of these fields, or more specifically, the ridge intersections, were hypothesized to indicate regions for enhanced severe convection. In addition, the sharpness of these ridge axes may correlate to an increase in severe potential. To test this theory, a mathematical formula was devised to quantitatively assess ridge strength. Then, using theta-e, lapse rate and their respective ridge strengths as predictors, a linear discriminate analysis was performed on dependent and independent datasets from the spring of 2008. Severe probabilistic forecasts were produced using the discriminate analysis and verified using two independent methods. Skill metrics calculated for the forecasts determined there is significant positive skill of the forecast technique. As a side note, this research attempted to determine the wavelength features (meso-α to meso-β) that recorded the highest level of skill using this method. No significant difference was noted in the different wavelength feature forecasts, possibly due to the verification method.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ACC:	Air Combat Command
AFWA:	Air Force Weather Agency
AOR:	Area of Responsibility
APE:	Available Potential Energy
BC:	Boundary Condition
BMJ:	Betts-Miller-Janjic
BRN:	Bulk Richardson Number
BS:	Brier Score
BSS:	Brier Skill Score
CAA:	Cold Air Advection
CAPE:	Convective Available Potential Energy
CI:	Confidence Interval
CONUS:	Continental United States
CPTP:	Cloud Physics Thunder Parameter
DMC:	Deep Moist Convection
DoD:	Department of Defense
IC:	Initial Condition
LFC:	Level of Free Convection
LLJ:	Low-Level Jet
LR:	Lapse Rate
MCC:	Mesoscale Convective Complex
MCS:	Mesoscale Convective System
MOS:	Model Output Statistics
MU:	Most Unstable
MUCAPE:	Most Unstable CAPE
NAM-WRF:	North American Mesoscale – Weather and Research Forecasting
NCEP:	National Center for Environmental Prediction
NSSFC:	National Severe Storms Forecast Center
NWP:	Numerical Weather Prediction

NWS:	National Weather Service
ORF:	Observed Relative Frequency
OWS:	Operational Weather Squadron
PBL:	Planetary Boundary Layer
RAOB:	Rawinsonde Observation
ROC:	Relative Operating Characteristic
ROCSS:	Relative Operating Characteristic Skill Score
RUC:	Rapid Update Cycle
SHI:	Severe Hail Index
SPC:	Storm Prediction Center
STR:	Severe Threat Region
TELR:	'Theta-e and Lapse Rate' forecast that uses theta-e, LR, theta-e ridge magnitude and LR ridge magnitude as predictors in the linear MDA
$TELR^{+}$:	'Theta-e and Lapse Rate *plus*' forecast that uses theta-e, LR, theta-e ridge magnitude, LR ridge magnitude, MUCAPE and 0 – 6 km shear as predictors in the linear MDA
USAF:	United States Air Force
VS:	Value Score
WAA:	Warm Air Advection

ACKNOWLEDGMENTS

I would like to thank the many people involved from the creation of the idea to the completion of the research process. First, thanks to Robert Baker, Master Sergeant, USAF (ret.) and Charles Hoffman, Master Sergeant, USAF, for their idea of using theta-e and lapse rate ridge intersections as regions for enhanced severe convection. Next, thanks to Wendell Nuss, PhD, for his expertise in severe weather forecasting and guidance in the research process. Also, thanks for the countless hours of computer programming help and database of programs used for this research. Thanks to Maj. F. Anthony Eckel, PhD, for his expertise and guidance in statistical analysis. Also, his MATLAB analysis code was essential in the creation of the skill metrics used in the analysis. Thanks to Patt Harr, PhD, for his guidance and programming code regarding the linear discriminate analysis used to create the forecasts. Thanks to Bob Creasey for his help in gathering data for this research and his countless hours of programming help. Thanks to Mary Jordan for her many hours of help with MATLAB programming. Thanks to Benny Neta, PhD, for his help in deriving the ridge analysis formula. Thanks to David Bright, PhD, of the Storm Prediction Center, for his suggestions and recommendations for experimental design. Finally, I'd like to thank my wife, Mandy Oltmer, for her support in my research endeavors.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The 26th Operational Weather Squadron (OWS) located at Barksdale Air Force Base provides weather support for all Department of Defense (DoD) installations in the central and southeastern United States, an area prone to severe weather. The DoD has assets venerable to severe weather totaling over $100 billion and over 500,000 personnel in the 26th OWS' area of responsibility (AOR). The forecasters at the 26th OWS use a variety of tools provided by the Air Force Weather Agency (AFWA) and the Storm Prediction Center (SPC) to determine the daily threat of severe weather in the AOR. Although these tools are quite helpful, there is still much uncertainty regarding the actual probability of severe weather at each installation. Currently, Air Combat Command (ACC) allows the 26th OWS forecaster a maximum false alarm rate of 50% for severe weather warnings, which means protecting assets twice as often as needed and wasting millions of DoD dollars hangering aircraft needlessly. The main objective of this study is to devise a severe weather forecast tool to objectively define a probabilistic forecast of severe weather for individual installations. This forecast will enable each base commander in the AOR to make smarter decisions about protecting assets when the probability warrants action, potentially saving millions of dollars.

Skill in severe weather prediction has increased dramatically over the past few decades due to integration of severe storms research into operations, advancements of numerical weather prediction (NWP) models, and more recently, integration of ensembles into the operational forecast procedure. Products from the SPC combine highly-skilled meteorologists with many years experience in severe weather forecasting, the latest in NWP and ensemble guidance, and comprehensive observational networks to produce short-term (0 – 6-hr) as well as longer-term forecasts (up to 7 days) of hazardous weather in the Continental United States (CONUS). AFWA produces NWP and ensemble forecast products tailored to give the 26th OWS forecaster a general threat region for daily severe weather.

The research in this study focuses on improving understanding of preferred regions of severe convection by determining skill (if any) in applying commonly used severe weather forecasting parameters, theta-e and lapse rate (LR) in a different manner than used currently. A mathematical formula devised quantitatively defines ridge strength of a 2-dimensional field. A general methodology of understanding environments conducive to severe convection was presented by Doswell (1987) as ingredients-based forecasting. By co-locating the most unstable (MU) parcel in the lowest levels and the MU environment, one could determine potential regions where convective activity could be maximized locally. This is attempted by locating the ridge axes of the maximum theta-e in the low-levels and the deep-layer LR, and focusing on the intersections of these ridge axes. As model resolution increases, the mesoscale features in the theta-e (as well as LR) ridge axes become apparent and could potentially serve to identify mesoscale regions of interest.

This research uses the 12km North American Mesoscale (NAM) Weather Research and Forecast (WRF) model output interpolated onto the 32-km, 221 grid using simple bilinear interpolation (NCEP website, n.d.). The minimum wavelength feature resolved by this 32-km gridded output is approximately 160 km. This length scale is typically associated with the meso-β range (20 – 200 km) and attempts to resolve organized severe convective systems such as squall lines, mesocyclones, and multi-cell thunderstorm regions (Ray, 1986). Although the operational NAM-WRF cannot yet resolve severe weather associated with the meso-γ scale (2 – 20 km) (supercells, microbursts, etc.), there are fine scale details (gradients) in parameter fields that could potentially increase skill in forecasting the mesoscale severe convective environment.

Will this wavelength feature give insight to the severe convective environment at each DoD installation and would overall organized severe weather forecasting skill increase if we looked at shorter (or longer) wavelength features?

A. OBJECTIVES OF THIS RESEARCH

To address this question, this thesis had four primary objectives:

(1) Test validity of this severe forecasting method (i.e., intersections of theta-e and LR ridges). Do the meso-β wavelength features properly capture the severe signal or is there a different wavelength feature (i.e.., meso-α or meso-γ) the forecast method verifies better?

(2) Apply linear discriminate analysis to demonstrate the validity of the theory using theta-e, LR, and a quantitative measure of their ridge axes as predictors to create a probabilistic forecasts of severe weather. This analysis was only performed in regions where the most unstable convective available potential energy (MUCAPE) $\geq$ 500 J/kg, 0 – 6 km shear $\geq$ 30 kts and 3-hr convective precipitation $\geq$ 0.01 in. co-existed to focus on credible ridge intersections where the probability of severe > 0%.

(3) Use common statistical metrics such as Brier Skill Score (BSS), Relative Operating Characteristic Skill Score (ROCSS) and reliability diagrams to determine skill in the method. Also, the Value Score (VS) will be assessed in an attempt to prove potential value to forecast users.

(4) Refine and test the method by adding two more predictors (MUCAPE and 0 – 6 km shear) to the discriminate analysis in attempt to increase forecast skill. The two forecasts (four and six predictors) will be compared for the validity in method and significant difference between the two.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND

A. INTRODUCTION

As mentioned in Chapter I, this research will assess severe weather potential using the intersections of the low-level theta-e and surface – 500 mb LR ridges. These two parameters were chosen in accordance with Doswell (1987), who wrote that severe storms are associated with deep moist convection (DMC), and that the three necessary ingredients are (a) a moist layer of sufficient depth in the low or mid-troposphere, (b) a steep enough lapse rate to allow for a substantial 'positive area,' and (c) sufficient lifting of a parcel from the moist layer to allow it to reach its level of free convection (LFC). The first two ingredients will be represented by the theta-e and LR ridge axes. These ridge axes indicate the local maximum of potential and conditional instability, respectfully. The third ingredient, initiation of convection, will be determined by the Betts-Miller-Janjic' (BMJ) convective parameterization scheme embedded in the NAM-WRF model. Although non-sheared environments can produce severe weather, it is well understood that vertical wind shear plays a critical role for thunderstorm organization which influences severity (see Rotunno & Klemp, 1982; Brooks, 1993; and Davies & Johns, 1993). For this reason, a minimum threshold of 0 – 6 km shear has been added as an ingredient of the severe recipe. Finally, to ensure the region of focus has severe potential from an instability perspective, a minimum threshold of MUCAPE $\geq$ 500 J/kg is required. This severe threat region (referred to from here on as STR) is defined by model-calculated most unstable (MU parcel from the surface to 255mb above the surface) CAPE (MUCAPE) $\geq$ 500 J/kg, 0 – 6 km shear $\geq$ 30 kts, and 3-hr convective precipitation $\geq$ 0.01 in. The following is a further discussion of the ingredients and their importance to severe weather.

B. INSTABILITY

There are many definitions of instability as noted by Schultz et al. (2000), but the key meaning is a positive amount of CAPE. Doswell (2001) wrote:

> Actual parcel instability leading to DMC is primarily associated with finite vertical displacements; hence, the key to the possibility for growth of convective storms is the presence of CAPE, not the environmental lapse rates alone (see Sherwood, 2000; Schultz et al., 2000). Not all situations with conditional instability are characterized by parcels with CAPE. Thus, the moisture content of the air is critical to knowing whether conditional instability actually contains the potential for parcels to become buoyant (i.e., to have CAPE). (p. 3)

This research uses the LR between the surface and 500 mb to simulate the environmental LR. The low-level temperature used in the LR calculation is the parcel 50 mb above the surface to achieve a diurnally independent measurement by avoiding the potential nocturnal temperature inversion at the surface. This use of LR as an instability qualifies as 'conditional' instability if the LR lies between the dry and moist adiabatic LRs (Shultz et al., 2000). This study will examine the LR 'ridge' to maximize the conditional instability and make some assumptions on the low- to mid-level advection characteristics. For usual eastward moving baroclinic systems in the CONUS, the longwave LR ridge typically protrudes northward for low-levels (depending on cloud cover) and eastward for mid-levels (i.e., approaching shortwave trough). These protrusions correspond to the warm and cold air advections (WAA/CAA) at their respective levels shown in Figure 1 (a) and (b). The LR ridge from the surface to 500 mb represents the vector between the maximum WAA/CAA in the low/mid-levels (shown in Figure 1 (c). As mentioned by Doswell (2001), the moisture content of the parcel also needs to be assessed in order to determine buoyancy.

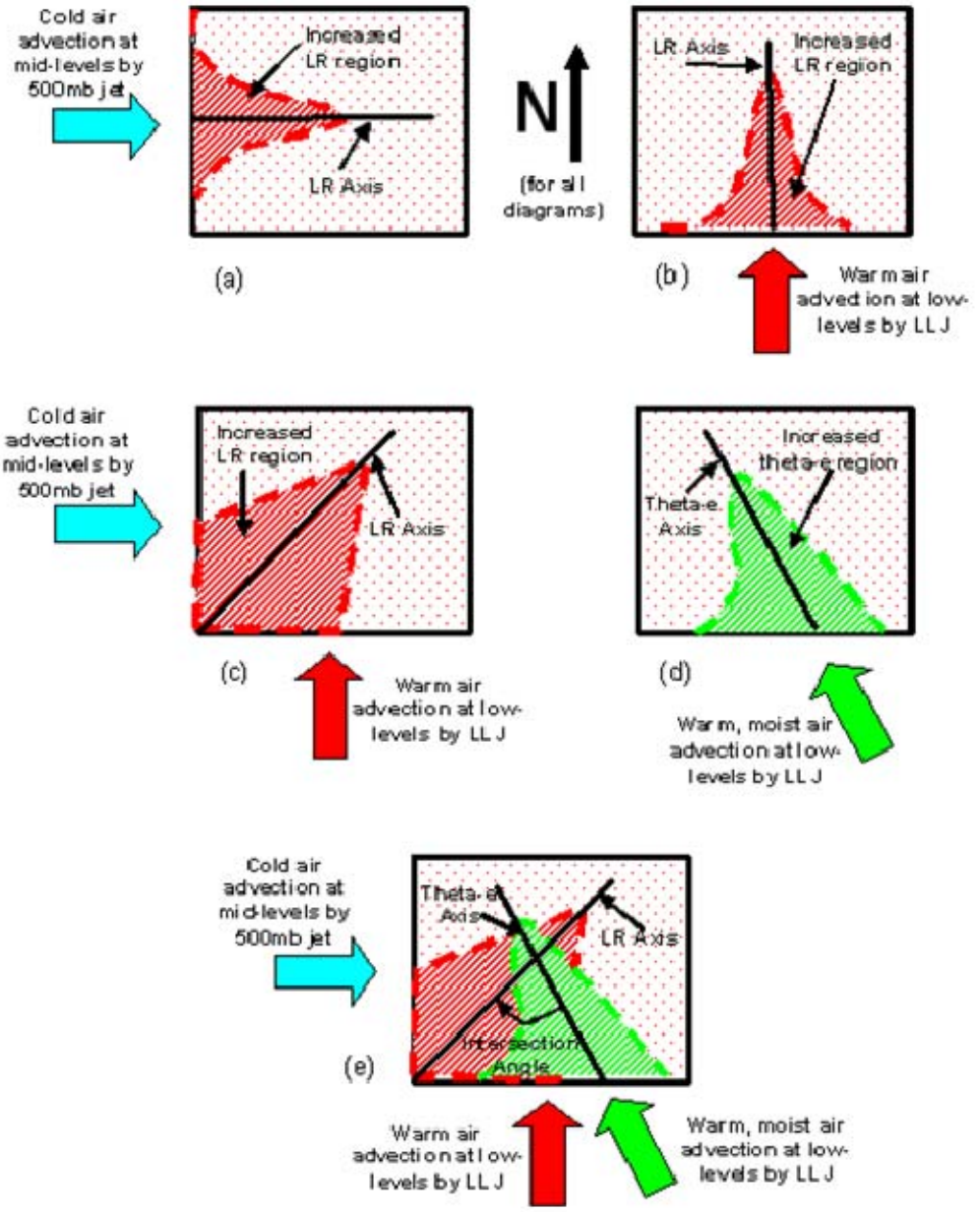


Figure 1. LR Ridge analysis for (a) CAA in the mid-levels by the 500mb jet, (b) WAA in the low-levels by the low-level jet (LLJ), (c) a combination of low- and mid-level WAA/CAA, (d) Theta-e ridge analysis for warm, moist air advection in the low-levels by the LLJ, and (e) the combination of the LR and Theta-e ridges to form the intersection angle.

The instability parameter used to define the STR is MUCAPE ≥ 500 J/kg. This threshold was defined somewhat arbitrarily due to forecasting experience. Bright (2003) developed the cloud physics thunder parameter (CPTP) for the Storm Prediction Center (SPC) and estimated convection required at least 100 – 200 J/kg of convective available potential energy (CAPE) to produce lightning. CAPE values on the order of 1000 – 3000 J/kg are typically associated with spring severe convection east of the Rockies. A low threshold of 500 J/kg narrows the domain considerably to define the STR and eliminates regions where severe thunderstorms have nearly zero probability of occurrence.

C. MOISTURE

East of the Rockies, the main source of low-level moisture for instability is the Gulf of Mexico. Synoptic-scale moisture flux is rapidly transported northward through the formation of a low-level jet (LLJ) in advance of next approaching upper-level trough from the west (Igau & Nielsen-Gammon, 1998). This low-level moisture advection is essential in diagnosing the actual available potential energy (APE) of the lifted parcel in the region of convection. McNulty (1995) noted the National Weather Service (NWS) forecasters typically analyze the theta-e ridges for areas of relatively deep moisture and unstable air needed for thunderstorm development. This northward advection of high theta-e air typically coincides with eastward advection of cool air aloft in typical severe weather producing baroclinic systems east of the Rockies. This northward advection is also referred to as a 'moist tongue' by Miller (1972) as a favorable region for severe convection.

Low-level theta-e can also be defined in terms of instability and related to a wide range of severe storms. Potential instability is defined as the decrease of theta-e with height by Rossby (1932). This warm, moist low-level air (i.e., high potential instability) in conjunction with steep LRs (i.e., high conditional instability) gives rise to large values of CAPE (i.e., high latent instability) as stated by Shultz et al. (2000). If the low-level potential instability were maximized, one could determine which lifted parcel could realize the largest buoyancy, a large factor in determining hail size (Fawbush & Miller, 1954; Foster & Bates 1956; and Miller 1972). Maddox (1980) correlated the strong, low-

level inflow of moist, unstable air to the continued growth of mesoscale convective systems (MCSs), a well-known severe weather producer. Thompson et al. (1994) researched a severe weather outbreak in November 1988 and tracked high theta-e air into the severe storm region in the central plains from the Gulf of Mexico. He determined significant increases in APE for deep convection were due to the low-level theta-e advection northward into the warm sector of the baroclinic system. Brooks et al. (1994), researched the development and maintenance of low-level mesocyclones in supercells and determined that increasing the low-level moisture content should affect the baroclinic generation of vorticity and might be an important variable in low-level meso-cyclogenesis. Multiple sources of research indicate the theta-e ridge to be crucial to the development of severe thunderstorms.

This research calculates the maximum theta-e from the surface to 700 mb to reveal the actual horizontal location of the maximum theta-e ridge, independent of height. This approach then allows for assessing regions of elevated convection.

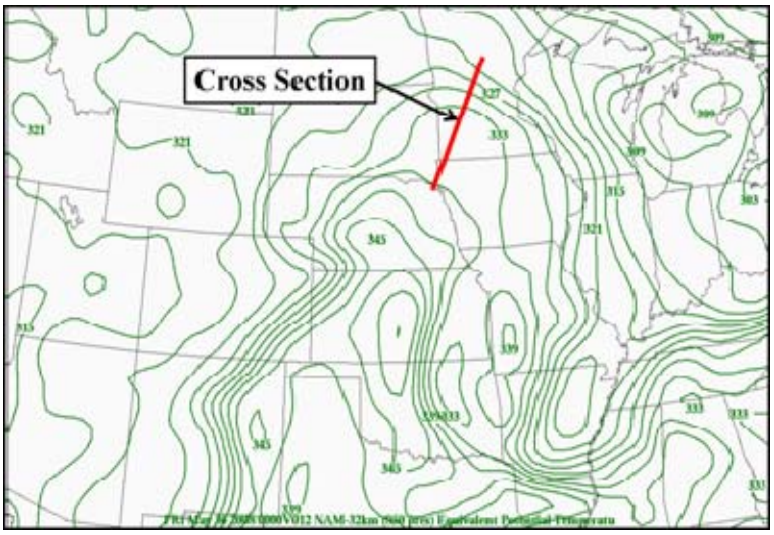


Figure 2. 071014/12Z Model 12-hr Forecast Valid at 071015/00Z – 850 mb Theta-e (K) (green contours) with Cross-Section line drawn (red line).

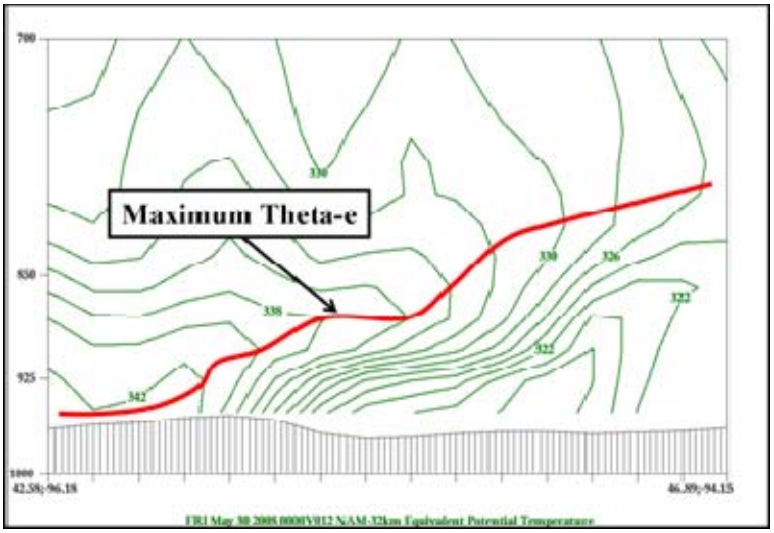


Figure 3. 071014/12Z Model 12-hr Forecast Valid at 071015/00Z – 1000 – 700 mb Theta-e Cross-Section (red line on Figure 2).

D. LIFTING MECHANISM

There are many ways to initiate surface-based convection by lifting planetary boundary layer (PBL) air to its LFC. These include mechanisms such as frontal boundaries, drylines, outflow boundaries, low-level convergence or sufficient surface heating. This research will not attempt to forecast convective initiation, but instead look in the area of model forecasted convective precipitation and delineate severe or no severe. As mentioned earlier, the NAM-WRF employs the BMJ convective parameterization scheme (Janjic, 1994). The BMJ parameterization initiates convection through the occurrence of three triggers, one of which is a potential limitation for thermodynamic profiles conducive to explosive severe convection (‘loaded-gun’ soundings) (COMET Program, 2006). Although this limitation exists, the scheme indicates regions of atmospheric overturning (of a certain depth), which captures the convective initiation ingredient of DMC.

E. VERTICAL SHEAR OF THE HORIZONTAL WIND

There are many different storm types, or convective modes that produce severe weather such as squall lines, supercells, pulse thunderstorms, MCCs, etc. It is widely understood that vertical wind shear structure is important for the thunderstorm organization which aids in the production of severe weather (Doswell & Bosart, 2001). The following are forecasting techniques used by the USAF from Miller (1972) concerning vertical wind shear and severe weather:

> (1) Pronounced veering (30 degrees or greater) between the low and middle-level winds is highly desirable for the development of severe storms. It appears the greater the directional shear between these levels, the more certain the development of severe activity and the more likely it subsequent movement and maintenance of intensity.
>
> (2) Speed differential is also desirable. In cases where the amount of veering between low- and mid-level flow is less than 30 degrees the middle-level must exceed the low-level winds by 30%. Pronounced veering accompanied by significantly stronger winds in the middle levels is the most favorable wind structure for tornado development.
>
> (3) The intersection of the middle- or upper-level jets and the horizontal shear associated with the low-level jet is a preferred zone for the development of mesosystems. (p. 70)

Miller (1972) addressed speed and directional vertical shear of the horizontal wind and its relation to severe convection. Veering wind with height (1) and jet intersection (3) can potentially be assessed by future research, but this research will focus on speed shear. More research has been conducted on the link between vertical wind shear and severe events such as supercells, MCCs, etc.

Supercells are characterized by rotating and tilted updrafts that typically produce severe convection due to the separated updraft/downdraft regions of the storm. Supercells produce all types of severe weather (Browning, 1964) including wind, hail, and tornadoes. Doswell (2001) notes that only 10% of organized DMC are supercells, but due to their intense updrafts, they create a disproportionate share of most forms of severe convection. Wilhelmson and Klemp (1978) suggested vertical wind shear is

confined to the lower portion (0 – 6 km) of the troposphere for development of supercells. Weisman and Klemp (1982) used a 3-D convective cloud model to investigate the relationship between buoyancy and unidirectional wind shear for storm types. Using the Bulk Richardson Number (BRN), a non-dimensional convective parameter, 0 – 6 km shear between 12 – 25 m/s (23 – 49 kts) is sufficient enough to produce a supercell with an environmental instability of 3000 J/kg.

A Mesoscale Convective Complex (MCC) is a large (> 100,000 km^2), long lived (> 6 hrs) storm that produces all types of severe weather (Maddox, 1980). Jirak and Cotton (2007) stated that shear parameters tend to be the best at distinguishing between MCS environments and environments that do not support organized convection. They found the mean 0 – 6 km shear to be 16.5 m/s (~ 32 kts) for MCS development.

Pulse-type thunderstorms that produce hail, wet microbursts, etc., typically exist in shear-free environments. As a result, this type of severe convection has a co-located updraft/downdraft, making them shorter lived and less organized than supercellular, squall line and MCC or MCS severe convection. Atkins and Wakimoto (1991) researched wet microbursts and their correlating environmental parameters. They found the typical 0 – 6 km wind shear to be approximately 4.4 m/s (~ 9 kts) for active microburst days.

The shear parameter used to define the STR in this research is 0 – 6 km shear > 30 kts. This threshold was shown to support the maintenance of organized severe convection such as supercells and MCSs. A weakness of this method by requiring vertical wind shear as an ingredient in severe weather forecasting is the non-forecast of ‘un-organized’ severe events such as wet microbursts (Wakimoto, 1991), convection with extreme amounts of CAPE and low shear (Johns et al., 1993), and unorganized severe convection. These severe events lack the requirement of vertical wind shear, and in some instances, call for little to no vertical wind shear.

III. DATA AND METHODOLOGY

A. INTRODUCTION

There are many steps involved in creating and verifying the validity of a new probabilistic severe thunderstorm forecast. First, the maximum theta-e in the low-levels, deep-layer LR and their respective ridge axes are calculated using the NAM-WRF model. Next, a severe threat region (STR) is defined as MUCAPE $\geq$ 500 J/kg, 0 – 6 km shear $\geq$ 30 kts and 3-hr convective precipitation $\geq$ 0.01 in. and the theta-e and LR ridge intersections in the STR are determined. Using a dependent dataset, statistical information is compiled in order to train the linear discriminate analysis. Next, the linear discriminate analysis (using four predictors – theta-e, LR, and their ridge magnitudes) is performed on the independent dataset to produce a probabilistic forecast of severe weather for each gridpoint at each forecast tau (0 – 24 hr). A second linear discriminate analysis is also performed (using two more predictors — MUCAPE and 0 — 6 km shear) in an attempt to refine the previous forecast. Finally, using radar data and reports of severe weather, each probabilistic forecast is verified to determine validity in the forecast method.

The research process was dissected into three main sections: (1) compiling the dependent dataset information; (2) compiling the statistical information on the dependent dataset and making a probabilistic severe forecast for the independent dataset; and (3) performing statistical analysis on the probabilistic forecasts to determine skill and value.

1. Dependent and Independent Datasets

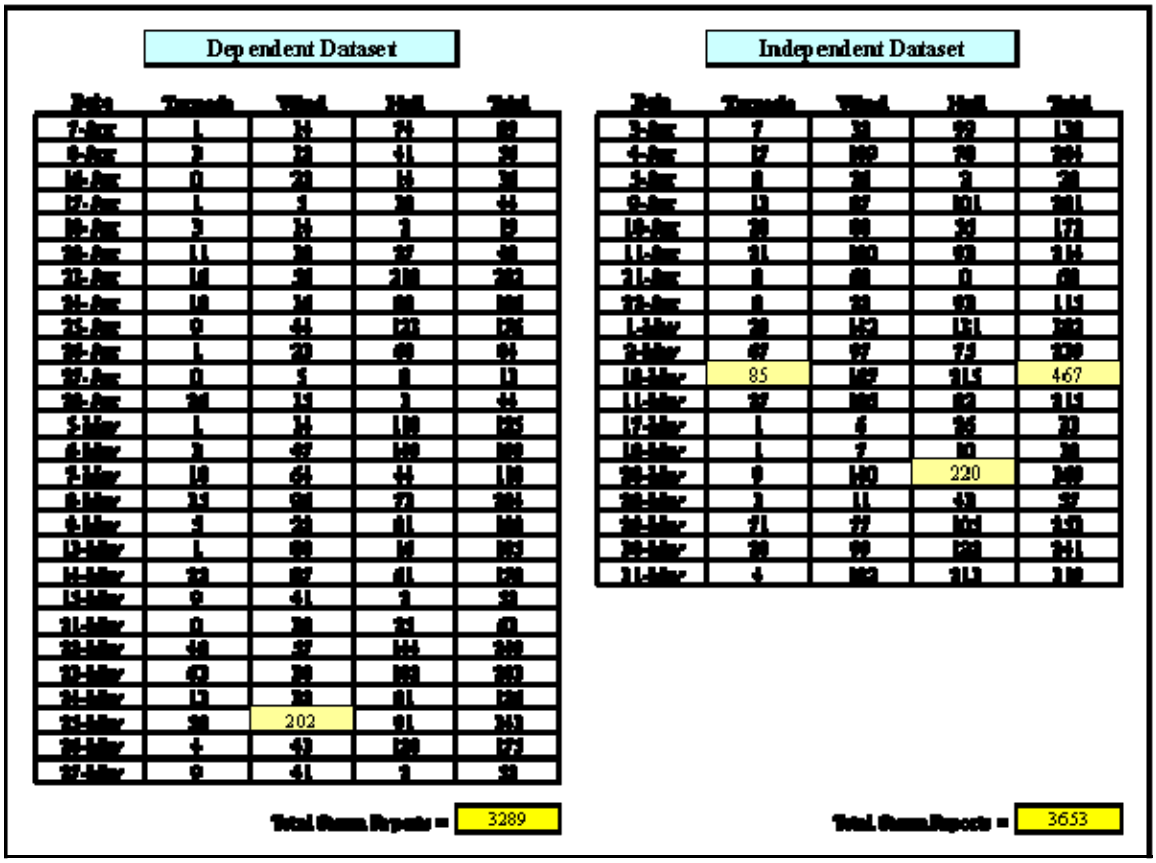
Since there are significant seasonal variances in the magnitudes of the forecast parameters, training over multiple seasons will lead to a lower predictability of severe weather. Due to this variability and the frequency of severe occurrence, a two-month period in the peak of severe weather season will be used for the dependent and independent datasets. The dependent and independent datasets are derived from severe storm events in the months of April and May of 2008. In an attempt to achieve

independence between the two datasets, subjective delineations were drawn between synoptic-scale storms over the two month period. There were 13 independent storm systems over the two-month period totaling 46 severe weather days. The dependent (independent) dataset will consist of the even (odd) numbered synoptic storms as depicted in Table 1. The SPC confirmed 3289 (3653) total tornado, hail and wind reports for the dependent (independent) dataset days as shown in Table 2. Although there are more severe weather days (8) in the dependent dataset, there are more severe reports (364) in the independent dataset.

Table 1. Subjective Division of Synoptic-Scale Storms between April and May 2008 with Independent Dataset hatched.

Storm	Date	Month
1	3, 4, 5	April
2	**7, 8**	**April**
3	9, 10, 11	April
4	**16, 17, 18, 20**	**April**
5	21, 22	April
6	**23, 24, 25, 26, 27, 28**	**April**
7	1, 2	May
8	**5, 6, 7, 8, 9**	**May**
9	10, 11	May
10	**13, 14, 15**	**May**
11	17, 18, 20	May
12	**21, 22, 23, 24, 25, 26, 27**	**May**
13	28, 29, 30, 31	May

Table 2. Dependent and Independent Dataset – Storm Prediction Center (SPC) Storm Reports from April – May 2008 with Individual Tornado, Wind, Hail and Total Report Maximums Highlighted.



2. Forecast Period

The morning (12Z) model run will be used to mimic daily severe weather operations for the following 24-hr period. Due to the non-linear nature of convection and the inability of current operational NWP to implicitly model these ageostrophic motions, model error growth decreases the limit of predictability for severe thunderstorm forecasting. Due to this constraint, the skill at the analysis (tau = 00) will determine the validity of the forecast technique. Assuming the technique has merit in identifying regions of severe weather (at the analysis), the following eight forecast taus (24 hrs) will be examined as well for skill.

B. STAGE I: COMPILE DEPENDENT DATASET

1. Develop FORTRAN Code to Develop Dependent Dataset Training

a. Read NAM-WRF Flat-files

The first task performed by the code is to read the operational 12-km NAM-WRF flat files which contain temperature, dew point, surface pressure, height, MUCAPE, 3 hr convective precipitation and a land/water differentiator. As mentioned before, the 12-km NAM-WRF is run for North America and is interpolated using simple bilinear interpolation to the Eta-221 32-km grid for ease of data transmission and storage. Wavelengths shorter than approximately 160 km (five gridspaces) are filtered by the transformation to a larger grid. At a horizontal gridspacing of 32 km, the NAM-WRF model domain has 277 zonal and 349 meridonal gridpoints with forty-three vertical levels.

b. Reduce Domain for Parameter Calculation

The domain was decreased from the full North American 277 × 349 grid to sub-domain of 81 × 81 gridpoints (shown in Figure 4). The smaller domain reduces algorithm computation time and focuses on the climatologically highest probability region for severe weather occurrence. The domain coordinates changed from the entire North American domain to a latitude of 26.32° to 43.16° and a longitude of –105.3° to -71.7° (CONUS east of the Rocky Mountains).

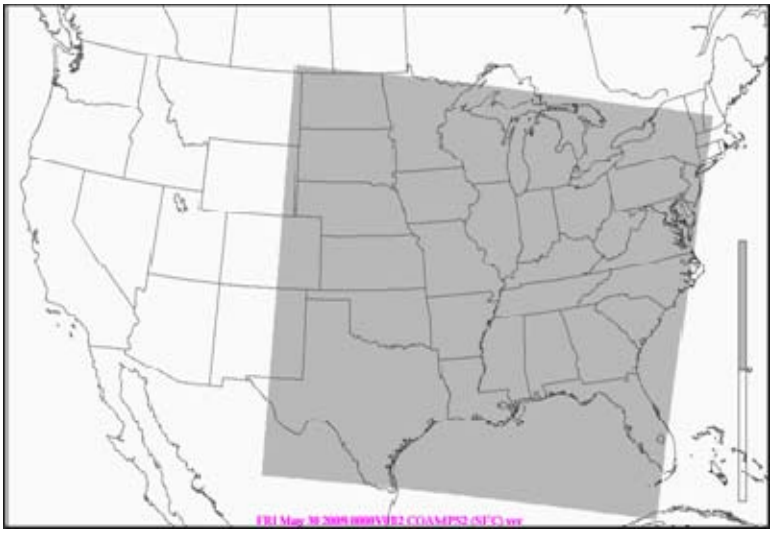


Figure 4. Reduced 81 × 81 Domain East of the Rockies (shaded).

c. Define STR in Model Space

As mentioned previously, a Severe Threat Region (STR) is defined by minimum thresholds of parameters that represent instability, shear, and convective precipitation in order to indicate general regions of severe weather potential. These thresholds were programmed using FORTRAN code to essentially act as a filter to exclude regions with near zero probability of severe convection. These thresholds are MUCAPE magnitudes ≥ 500 J/kg (Figure 5), 0 – 6 km shear ≥ 30 kts (Figure 6), and model forecast 6-hr cumulative convective precipitation (valid through three hours before and after tau) ≥ 0.01 in (Figure 9). The 3-hr cumulative convective precipitation fields valid before and after (Figure 7 and Figure 8, respectively) are combined to create the regions (Figure 9) in which convection is forecast to occur in the 6-hr window. The purpose of using this window instead of the 3-hr cumulative precipitation ending at tau was to potentially account for spatial and temporal forecast error (model placement of squall line, timing of initiation of convection, etc.). This 6-hr window of convective precipitation could act as a source of error in the delineation of the STR. The objective of the forecast is to produce an instantaneous probability of severe weather at the forecast tau, but the six-hour window will largely over estimate the region in which severe

convection has potential (STR) at a particular forecast tau. In the future, the instantaneous convective rain-rate could be used in place of the 6-hr convective precipitation. The overlap of all three parameters shown in Figures 5, 6, and 9 create the STR in Figure 10.

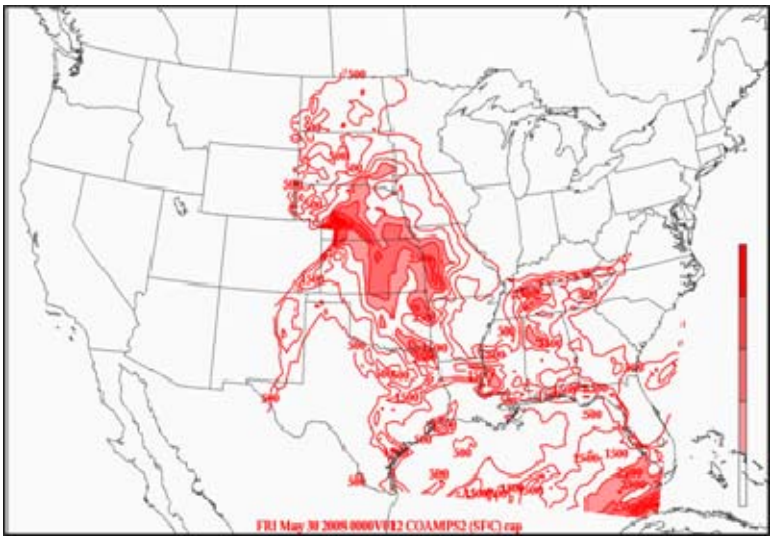


Figure 5. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z) MUCAPE (red lines with > 2000 J/kg highlighted).

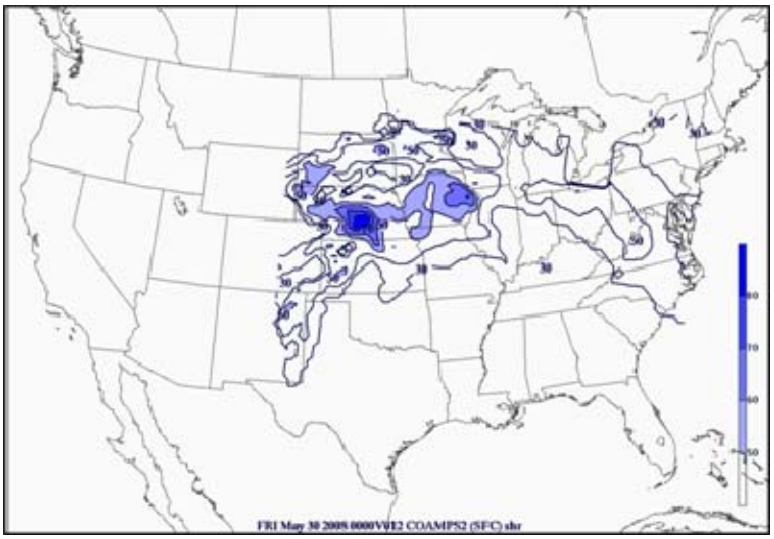


Figure 6. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z) 0 – 6 km Shear (blue lines with > 50 kts highlighted).

Figure 7. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 29/21Z – 30/00Z) 3-hr Cumulative Convective Precipitation (green contours).

Figure 8. 080529NAM-WRF 12Z Run – 15-hr Forecast (valid 30/00Z – 30/03Z) 3-hr Cumulative Convective Precipitation (green contours).

Figure 9. 080529NAM-WRF 12Z Run – Combination of the 12/15-hr Forecasts (valid 29/21Z – 30/03Z). 6-hr Cumulative Convective Precipitation (green contours).

Figure 10. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Severe Threat Region (STR) (black line).

d. Mathematically Define Ridge Analysis

In order to locate model forecast theta-e and lapse rate ridges, a mathematical formula needs to be defined to detect the ridge axis. This is accomplished by Equation (1), expanded in Equation (2).

$$\text{Ridge Axis} = -\nabla \bullet \left(\frac{\nabla \theta_E}{|\nabla \theta_E|} \right) \qquad (1)$$

$$\text{Ridge Axis} = -\frac{d}{dx}\left(\frac{\frac{d\theta_E}{dx}}{\sqrt{\left(\frac{d\theta_E}{dx}\right)^2 + \left(\frac{d\theta_E}{dy}\right)^2}} \right) - \frac{d}{dy}\left(\frac{\frac{d\theta_E}{dy}}{\sqrt{\left(\frac{d\theta_E}{dx}\right)^2 + \left(\frac{d\theta_E}{dy}\right)^2}} \right) \qquad (2)$$

First, the gradient (and magnitude of the gradient) of the two-dimensional parameter field is calculated. The gradient is a vector field indicating locations where the parameter changes with respect to x and y. The magnitude of this vector is a scalar with the same units as the gradient. Next, divide the gradient vector by its magnitude to determine the unit vector field perpendicular to the contour field as shown in Figure 12. The 850 mb theta-e contours (Figure 11) indicate the main theta-e ridge axis through central Oklahoma, Kansas and Nebraska with secondary ridge axes branching eastward toward southern Iowa and northern Missouri.

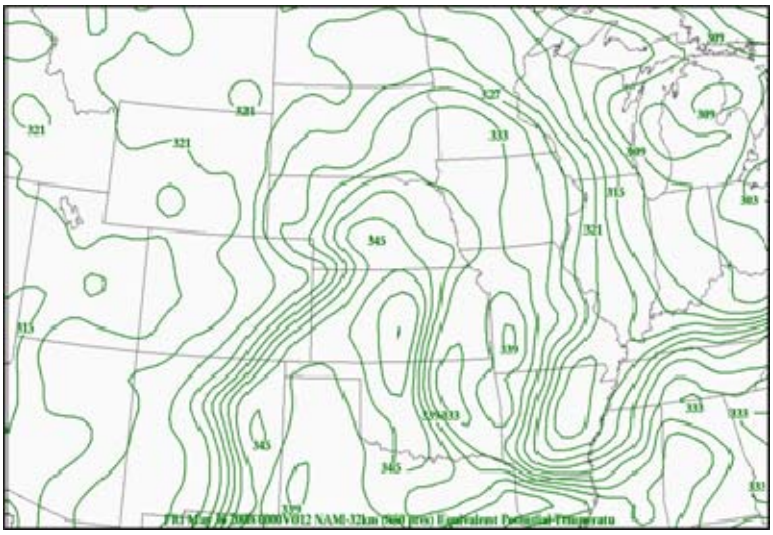


Figure 11. 050507/12Z Model 12-hr Forecast Valid at 050607/00Z – 850 mb Theta-e (green contours).

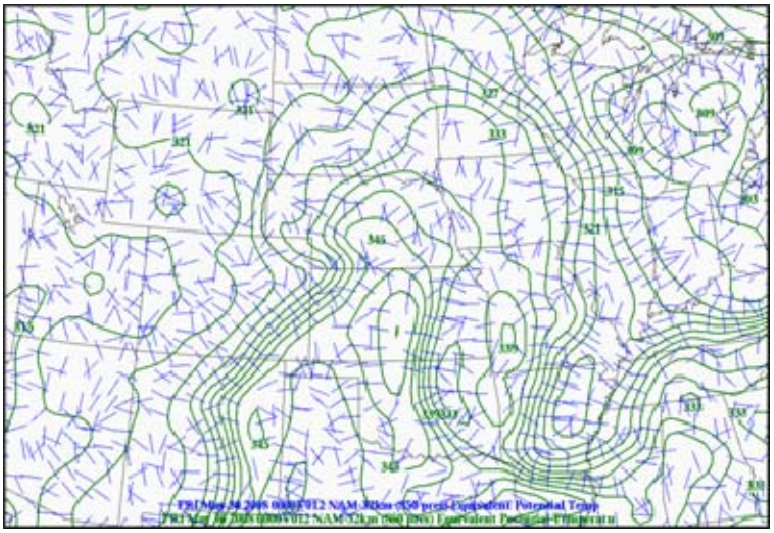


Figure 12. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Theta-e (green contours) with Unit Gradient Vector Field (blue arrows) Pointing Toward the Local Theta-e Maximum.

Lastly, calculate the convergence of this vector field to determine the placement of the ridge axes since the vectors point toward local maximums as shown in Figure 12. A quantitative measure of ridging, indicated by the contours in Figure 13,

corresponds to convergence values ranging between 0 and $-20E-5$, with stronger ridging coinciding with larger absolute values.

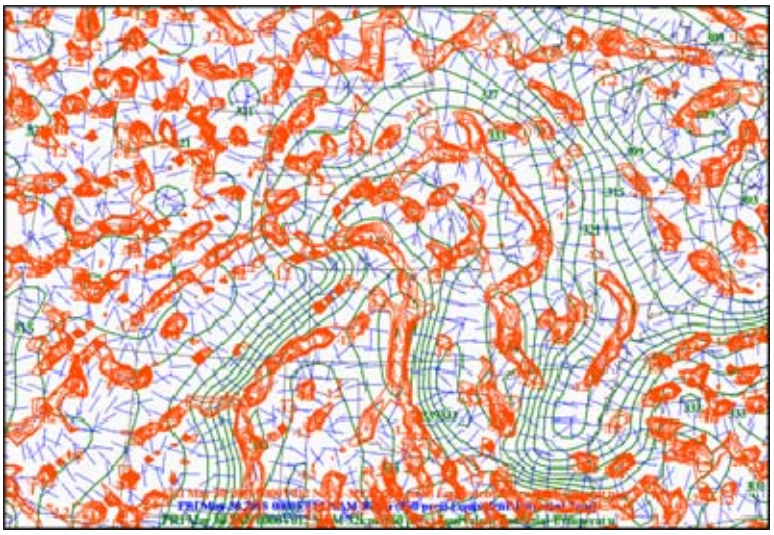


Figure 13. 050507/12Z Model 12-hr Forecast Valid at 050607/00Z – 850 mb Theta-e Field (green contours) with Unit Gradient Vector Field (blue arrows) Pointing Toward the Local Theta-e Maximum and the Ridge Axis Defined by Convergence of the Unit Gradient Vector Field (orange contours).

e. Theta-e Ridge Calculation

A subroutine was written to calculate the maximum theta-e value for each gridpoint in the lowest fourteen model pressure levels (surface to 700 mb) (shown in Figure 14). The ridge analysis described above was performed on this two-dimensional field to determine the ridge axes of the maximum low-level theta-e (Figure 15). Each model gridpoint has an assigned value of maximum theta-e and theta-e ridge magnitude.

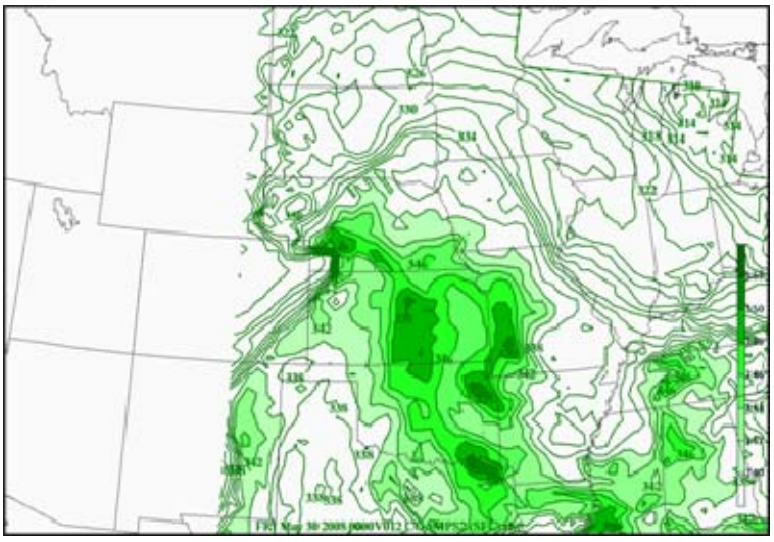


Figure 14. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Maximum Theta-e (Surface – 700mb) (green contours with > 340 K highlighted).

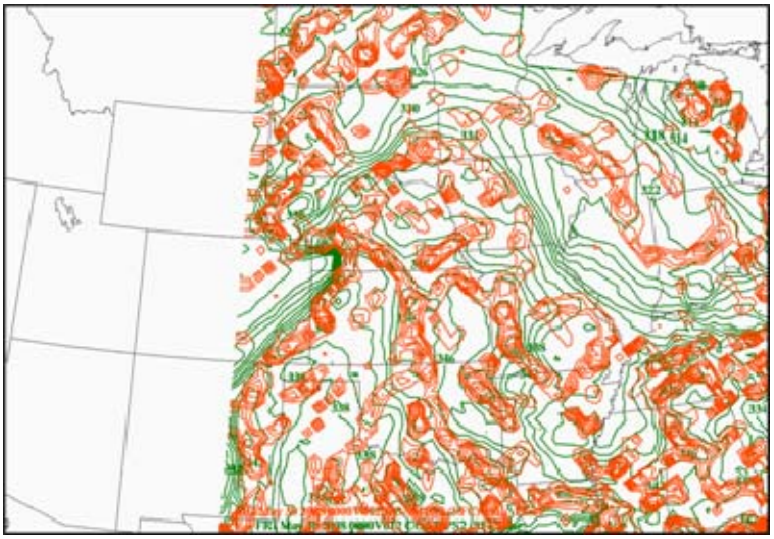


Figure 15. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Maximum Theta-e (Surface – 700mb) (green contours) and Ridges (orange contours).

f. LR Ridge Calculation

The LR ridge calculation is much like the theta-e ridge calculation, but the approximate LR between the surface and 500 mb is first computed. A lapse rate is defined as a temperature difference between two vertical levels divided by its depth. The FORTRAN code locates and assigns the level 50 mb above the model surface as the lower threshold used in the temperature difference to avoid potential surface temperature inversions. The height (above ground) of this level is also determined for the depth calculation in the denominator. Upon determining the two-dimensional LR field (Figure 16), the ridge analysis is performed (Figure 17). Once again, each model gridpoint is assigned a value of LR and LR ridge magnitude.

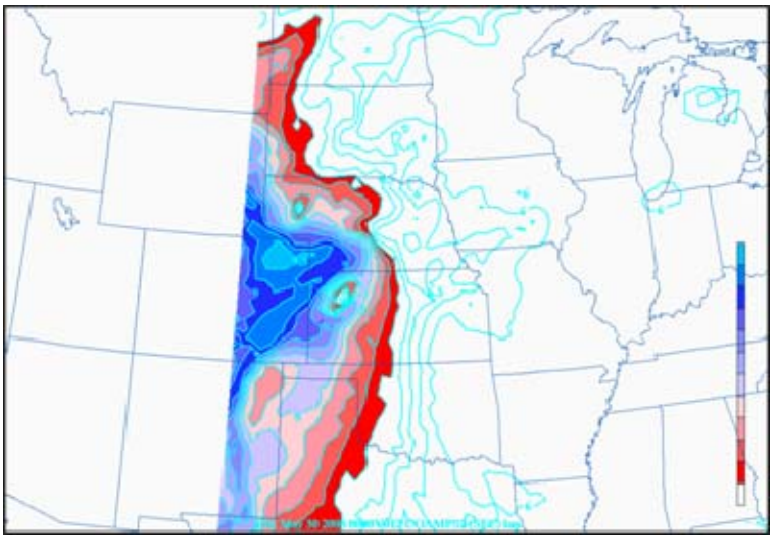


Figure 16. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Lapse Rate (LR) (Surface – 500mb) (blue contours with > 7 deg C/km highlighted).

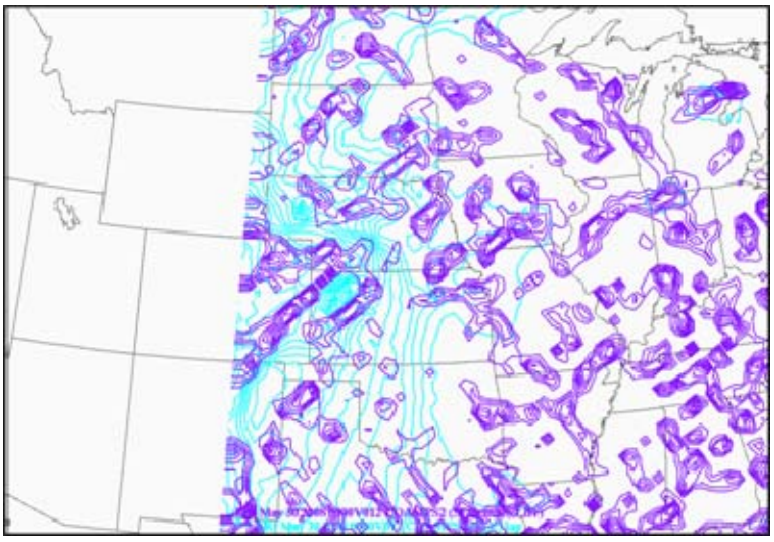


Figure 17. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Lapse Rate (LR) (Surface – 500mb) (blue contours) and Ridges (purple contours).

g. Training with the Dependent Dataset

The analysis and first eight forecast taus (00 – 24-hr) of each dependent dataset model run are input into this process described above. The code will first define a STR for each forecast tau. Next, it will determine if the theta-e and LR ridges axes intersect inside the STR by checking for co-existing values of positive convergence (positive convergence defined as ridge axis). Where all of the above are true, the model gridpoint is designated as a 'yes' severe forecast. All gridpoints where no intersection occurs (even inside the STR) are designated as 'no' severe forecasts.

h. Verification

A combination of two independent methods was used to verify the severe weather forecasts. The first verifies severe weather by storm reports disseminated by the SPC. These are reports of tornadoes, hail greater than 0.75 in. and damaging winds relayed from the public, storm spotters and local officials through regional NWS offices. Each storm report has a latitude and longitude location identifier which verifies the gridpoint corresponding 32 × 32 km gridbox in which it falls (shown in Figure 18). Each

gridpoint has a logical variable, true (1) or false (0), for severe verification. The second method to verify severe occurrence utilizes mosaic composite radar reflectivities ≥ 50 dBz (Figure 19). Each gridbox will be analyzed over a ± 15 min. window around the forecast tau to simulate the instantaneous convection. The highest reflectivity (dBz) measured in each 32 × 32 km gridbox over the 30 min. period will be assigned to the corresponding gridpoint to simulate the most severe reflectivity at the specific forecast tau. The reflectivity threshold of 50 dBz was chosen based on the WSR-88D severe hail algorithm evaluated by Lenning et al. (1998). The severe hail index (SHI) in the algorithm uses reflectivity ≥ 50 dBz as an indicator for the presence of severe hail in a storm cell. The presence of 50 dBz reflectivity in a cell does not guarantee severe hail falling to the ground, but will act as an approximation in this research. Meeting the 50 dBz threshold at any location in the 32 × 32 km gridbox assigns true (1) verification for the corresponding gridpoint.

Figure 18. Storm Reports valid 080529/2345Z – 080530/0015Z (red contours).

Figure 19. Radar Reflectivities ≥ 50dbz valid 080529/2345Z – 080530/0015Z (red area).

Finally, in an attempt to verify only 'organized' severe convection, the verification procedure requires the forecast vertical shear to exceed the STR requirements (0 – 6 km shear ≥ 30kts). This requirement reassigns gridpoints with a severe report or radar reflectivities ≥ 50 dBz that fail to meet the vertical shear pre-requisite to have a false (0) value for severe verification. This requirement will allow the method to have proper statistical consideration for the verification of what the method is attempting to forecast. For example, the combination of the verification methods is shown below in Figure 20. The reflectivity ≥ 50 dBz in western North Dakota was removed from the final verification due to its vertical shear < 30 kts (vertical shear shown in Figure 6).

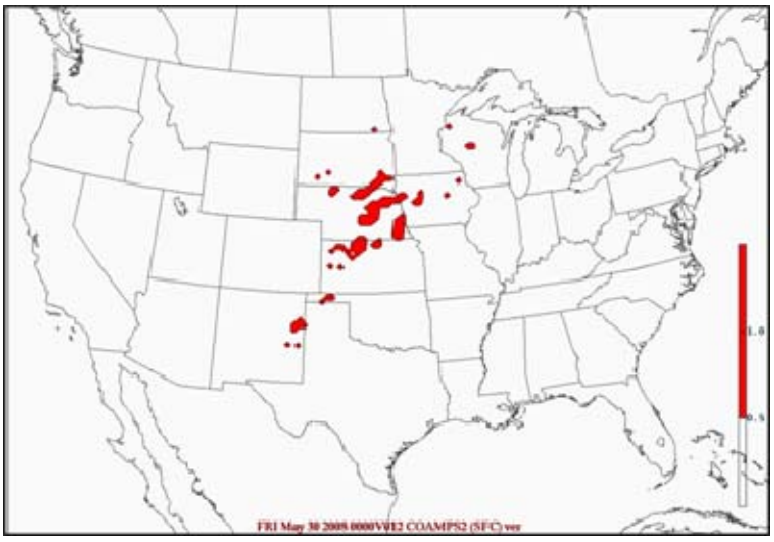


Figure 20. Verification of Severe Weather using Radar Reflectivity Composites and Storm Reports valid 080529/2345Z – 080530/0015Z (red area).

A 'yes' forecast gridpoint will be verified within a 96 × 96 km gridbox centered on the forecast gridpoint. This is the approximate area of a circle with radius of 34 mi centered on each gridpoint. The SPC verifies their severe weather in a similar fashion, but the radius is 25 mi. (SPC webpage, n.d.). The reason for expanding the verification of a forecast from the corresponding gridbox (32 × 32 km) to include the surrounding eight gridboxes is because the relatively low probability of severe weather at any given point. Also, this will give the model forecasts a reasonable chance to verify with small timing or placement errors.

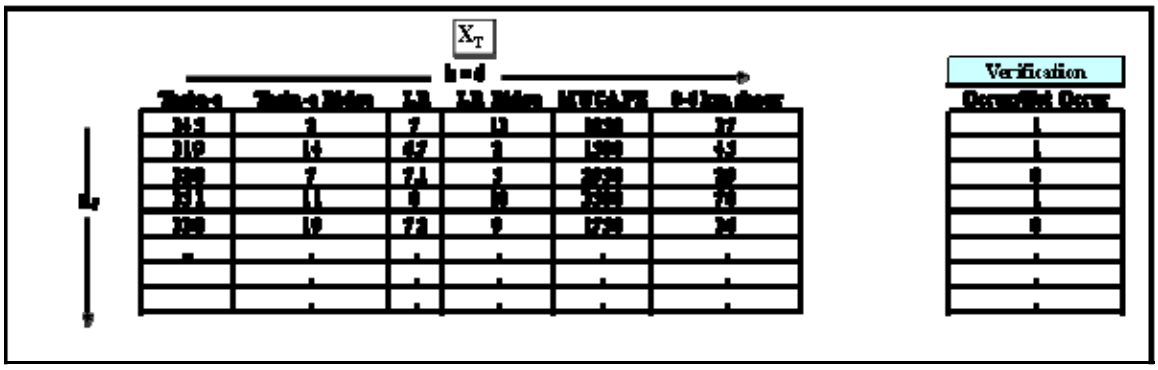
C. STAGE II: COMPILE STATISTICAL INFORMATION FROM THE DEPENDENT DATASET

1. Results of Compilation of Dependent Dataset

Upon completing Stage I, a forecast is made and verified for each gridpoint (6561), at each tau (9) of each model run in the dependent dataset (27). This gives a total of 1,594,323 gridpoint forecasts, but only the 'yes' forecasts are important for training. At this point it is necessary to compile data from all 'yes' forecast gridpoints, from all

taus, from all dependent dataset model runs. This compilation is essentially the gridpoints where the theta-e and LR ridge axes cross in the STR. At each gridpoint, there are corresponding parameter values that defined the STR such as MUCAPE and shear as well as values for theta-e and LR and their corresponding quantitative ridge strengths. The values of these parameters are used in the discriminate analysis probabilistic forecast in Stage III to determine forecasting skill. The output from Stage I can be visualized in Table 3.

Table 3. Example of Parameter Values for each Gridpoint for a 'Yes' Forecast, $\mathbf{X_T}$, and the Corresponding Verification for the Dependent Dataset.



This parameter matrix needs to be divided into two groups for statistical variable calculations. This is achieved by grouping like verifications of occurred and not occurred gridpoints. This can be portrayed in matrix format by:

$$X_T = x_1 + x_0 \tag{3}$$

X_T is the entire parameter matrix shown in Table 3 and has dimensions $(n_T \times k)$. Each sub-matrix, x_1 and x_0 correspond to the parameter matrices of the occurred and not occurred verification and have dimensions $(n_1 \times k)$ and $(n_0 \times k)$, respectively.

2. Probabilistic Forecast Theory

In Stage III, the probabilistic classification of the independent dataset forecasts into each group uses linear discriminate analysis. The use of discriminate analysis for forecasting was first applied in the meteorological community by Miller (1962) who forecasted airfield ceiling heights. McNulty (1981) also used four predictors to forecast thunder, severe thunder, or no thunder for the National Severe Storms Forecast Center (NSSFC) daily convective outlook.

In addition to the discriminate analysis, and an application Bayes' Theorem (Wilks, 2006) will allow us to create a probabilistic forecast using:

:

$$\Pr\left\{\text{Group g}\middle|x_f\right\} = \frac{p_g\left(\left|\left[S_g\right]\right|^{-1/2}\exp\left[-\frac{1}{2}\left(x_f-\overline{x}_g\right)^T\left[S_g\right]^{-1}\left(x_f-\overline{x}_g\right)\right]\right)}{\sum_{h=1}^{G}p_h\left(\left|\left[S_h\right]\right|^{-1/2}\exp\left[-\frac{1}{2}\left(x_f-\overline{x}_h\right)^T\left[S_h\right]^{-1}\left(x_f-\overline{x}_h\right)\right]\right)} \qquad (4)$$

Group g in Equation (4) specifies classification into a specific group (group 1 = severe and group 0 = no severe). The variable x_f represents the forecast parameters (theta-e, LR, ridge magnitudes, MUCAPE and shear) in the independent dataset 'yes' forecasts. Equation (4) states, given a 'yes' forecast for a gridpoint in the independent dataset (intersection of the theta-e and LR ridges in the STR – denoted as x_f), what is the probability of belonging to group g. The variable p_g refers to the prior probability (dependent dataset) of group g membership. For example, p_1 is basically the relative frequency of a verified 'yes' forecast using the training dataset. For the counter, p_0, this is relative frequency of the unverified 'yes' forecast.

$$p_1 = \frac{n_1}{n_T} \qquad (5)$$

$$p_0 = \frac{n_0}{n_T} \tag{6}$$

$\left[S_g\right]$ refers to the sample covariance matrix for group g. To simplify Equation (4), a pooled estimate of the covariance, $\left[S_{pool}\right]$, is calculated.

$$\left[S_{pool}\right] = \frac{n_1 - 1}{n_1 + n_0 - 2}\left[S_1\right] + \frac{n_0 - 1}{n_1 + n_0 - 2}\left[S_0\right] \tag{7}$$

The vector $\bar{x}_g$ in Equation (4) is a column vector $(k \times 1)$ that represents the sample mean of each dimension. Equation (4) now simplifies to Equation (8) to for the probability of severe weather:

$$\Pr\left\{\text{Group } 1 \middle| x_f\right\} = \frac{p_1\left(\exp\left[-\frac{1}{2}\left(x_f - \bar{x}_1\right)^T\left[S_{pool}\right]^{-1}\left(x_f - \bar{x}_1\right)\right]\right)}{p_0\left(\exp\left[-\frac{1}{2}\left(x_f - \bar{x}_0\right)^T\left[S_{pool}\right]^{-1}\left(x_f - \bar{x}_0\right)\right]\right) + p_1\left(\exp\left[-\frac{1}{2}\left(x_f - \bar{x}_1\right)^T\left[S_{pool}\right]^{-1}\left(x_f - \bar{x}_1\right)\right]\right)} \tag{8}$$

3. Statistical Variable Calculations

a. Determining Skill in Theta-e and LR Ridge (TELR) Intersection Forecast

In order to first determine if using the theta-e and LR ridge intersections have skill in forecasting severe weather, only the first four parameters of X_T (in Table 3), theta-e, LR, and their ridge magnitudes will be used to train the forecast. In this case, g = 1 and g = 0 remain the same as the occurred and not occurred groups; the total number of positive forecasts, as well as the number verified and not verified, n_T, n_1 and n_0 respectively, remain the same, and the only variable that changes is k from six

parameters to four. The new x_1, x_0, $\overline{x}_1$, $\overline{x}_0$, $[S_1]$, $[S_0]$, $\left[S_{pool}\right]$, p_1 and p_0 are calculated for insertion into Equation (4) to determine the probabilistic forecast (Stage III) for each independent 'yes' forecast (x_f). This forecast will be referred to as the 'theta-e and lapse rate' (TELR) forecast.

b. Determining an Increase in Skill Using All Six Parameters (TELR$^+$)

In an attempt to increase skill over the TELR forecast, MUCAPE and 0 – 6 km shear will be added to the four predictors used to train the severe weather forecast. For this calculation, $k = 6$ as previously shown in Table 3. The same statistical variables calculated for the TELR forecast (x_1, x_0, $\overline{x}_1$, $\overline{x}_0$, $[S_1]$, $[S_0]$, $\left[S_{pool}\right]$, p_1 and p_0) are recalculated using $k = 6$. This forecast will be referred to as the 'theta-e and LR plus' (TELR$^+$) forecast.

D. STAGE III: MAKE TELR AND TELR$^+$ FORECAST FOR INDEPENDENT DATASET USING STATISTICS FROM DEPENDENT DATASET

1. Repeat Steps a. – f. in Stage I on the Independent Dataset

Steps a. – f. in Stage I in Section B. performed on the independent dataset define the STR and determine the gridpoints with theta-e and LR ridge intersections for each tau (9) of each independent dataset model run. This produces a 'yes' or 'no' forecast (as in Stage I) for each gridpoint in the 81 x 81 domain. In order to create the probabilistic TELR and TELR$^+$ forecasts based on the trained dataset, the parameter matrices x_f (shown in Table 4 below where $k = 4$ and $k = 6$, respectively) are defined as the parameter values at each independent dataset forecast 'yes' gridpoint.

Table 4. Example of Parameter Values for each 'Yes' Forecast (x_f) in the Independent Dataset using k = 4 (TELR Forecast on left) and k = 6 (TELR$^+$ Forecast on right).

x_f for TELR Forecast

k = 4

Theta-e	Theta-e Ridge	LR	LR Ridge
345	2	7	13
319	14	6.7	2
328	7	7.1	5
351	11	8	18
338	19	7.2	9
..	.	.	.
	.	.	.
	.	.	.

x_f for TELR$^+$ Forecast

k = 6

Theta-e	Theta-e Ridge	LR	LR Ridge	MUCAPE	0-6 km shear
345	2	7	13	1850	37
319	14	6.7	2	1500	45
328	7	7.1	5	2950	29
351	11	8	18	3500	70
338	19	7.2	9	1750	36
..	.	.	.	.	.
	.	.	.	.	.
	.	.	.	.	.

2. Make and Verify Probabilistic Forecasts Using TELR and TELR$^+$

After creating the two matrices of x_f for the independent dataset, apply Equation (8) to determine the probability of severe weather at each gridpoint using both forecasts and verify each gridpoint (step h. in Stage I). Each gridpoint in the 81 × 81 domain, for each tau, for each model run now have (TELR and TELR$^+$) probability forecasts as well as an occur/not occurred verification. An important thing to note in the analysis of the dependent dataset is the probability of severe weather at 'non-intersections' inside the STR (shown in Table 5).

Table 5. Probability of Severe Occurrence (%) for Non-Intersections Inside the STR per Forecast Tau.

00	03	06	09	12	15	18	21	24
0.89726	2.5601	1.4206	1.5371	6.31	1.0111	1.2574	1.0873	1.1686

This tau dependent probability value is assigned to the 'non-intersections' inside the STR for the independent dataset forecast. By definition, the STR delineates the positive probability of organized severe weather, so the gridpoints outside the STR will be assigned a 0% forecast probability.

E. STAGE IV. STATISTICAL VERIFICATION OF PROBABILISTIC FORECAST

1. Dataset Dilution Problem

In an attempt to alleviate inflated statistical results, two sets of metrics will be calculated. The first skill calculation (industry standard) includes all gridpoints in the entire domain (6561), even regions where there is a nearly 0% probability of severe weather occurrence. This is referred to as dataset dilution by Eckel (2008) where inflating the verification dataset by including forecasts/verifications with nearly zero uncertainty (e.g., nearly 0% probability of severe weather in a region of strong mid-level ridging) artificially improves the skill metric. The second metric calculation will involve only grid points in the STR. This is defined as the region in which there is a greater than zero probability of organized severe weather occurrence. This shedding of 'correct rejections' (shown as **d** in the 2×2 contingency table, Figure 21) should make the skill metric more meaningful. A downfall to the second method will be regions outside the STR will be totally ignored as having potential for organized severe weather. If there is organized severe occurrence outside the STR this will not count negatively against the skill for the second metric calculation. For this research, both metrics will be displayed for comparison purposes to other research.

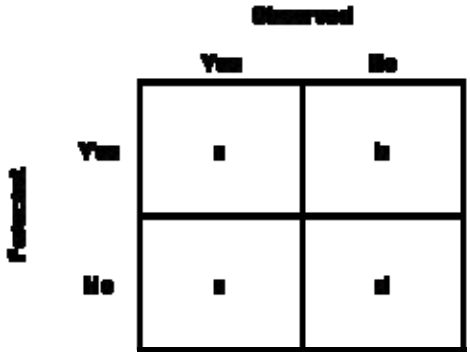


Figure 21. 2×2 Contingency Table.

2. Skill Metrics

In order to evaluate skill of the TELR and TELR^{+} forecasts, metrics such as Relative Operating Characteristic Skill Score (ROCSS) and the Brier Skill Score (BSS) are calculated. A second calculation of BSS can be calculated by its decomposed components (reliability, resolution and uncertainty) derived from the reliability diagram. This diagram also allows diagnosis of forecast biases that can potentially be alleviated by simple calibration methods. Finally, Value Score (VS) will be assessed to determine potential value to forecast users with different risk tolerances. The formulas for the metrics can be found in Appendix II and more information can be found in Wilks (1995).

IV. DATA ANALYSIS AND RESULTS

A. RESULTS

1. Graphical Example of Intersections, TELR and TELR^{+} Probabilistic Forecasts and Verification

The 300 mb, 500 mb, 700 mb, 850 mb, 925 mb, and the surface analysis (SPC webpage, n.d.) plots valid 12Z, 29 May 2008 are shown in Figure 22 – Figure 27 below to get a synoptic understanding of the independent dataset day example. The SPC 'Day 1 Convective Outlook' (valid 29/13Z – 30/12Z) and storm reports for the 24-hr period (29/12Z – 30/12Z) are shown in Figure 28 – Figure 29. This was one of the few 'high' risk severe outlooks issued by the SPC in 2008.

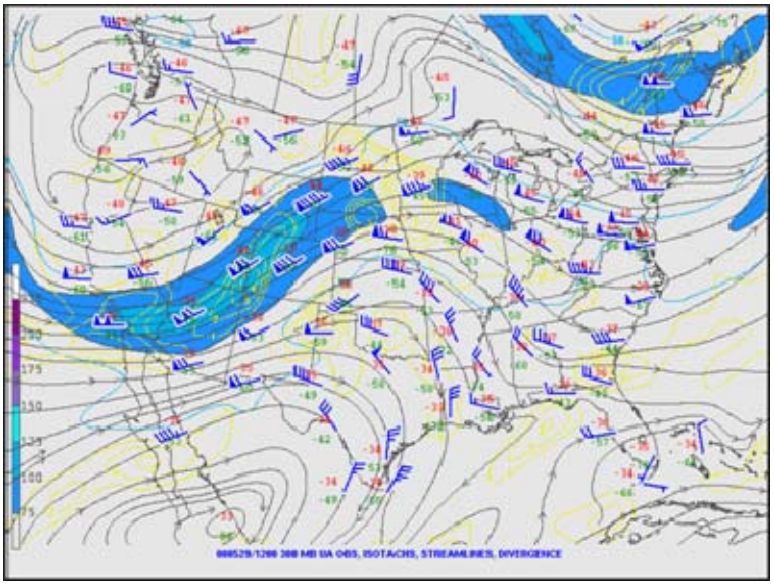


Figure 22. 080530/00Z 300 mb Analysis (From: SPC webpage, n.d.).

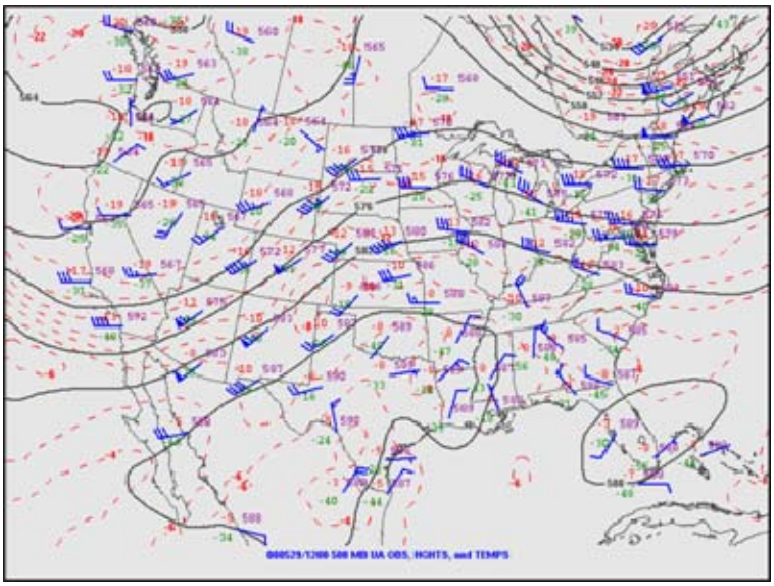


Figure 23. 080530/00Z 500 mb Analysis (From: SPC webpage, n.d.).

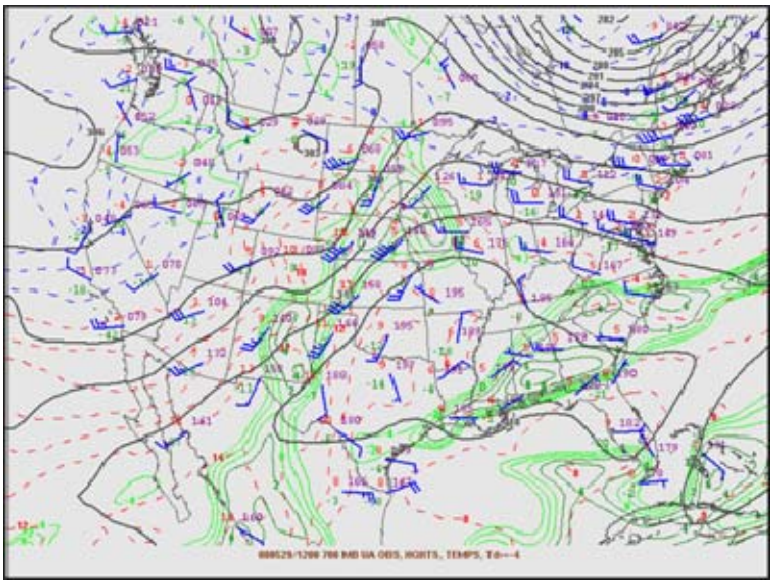


Figure 24. 080530/00Z 700 mb Analysis (From: SPC webpage, n.d.).

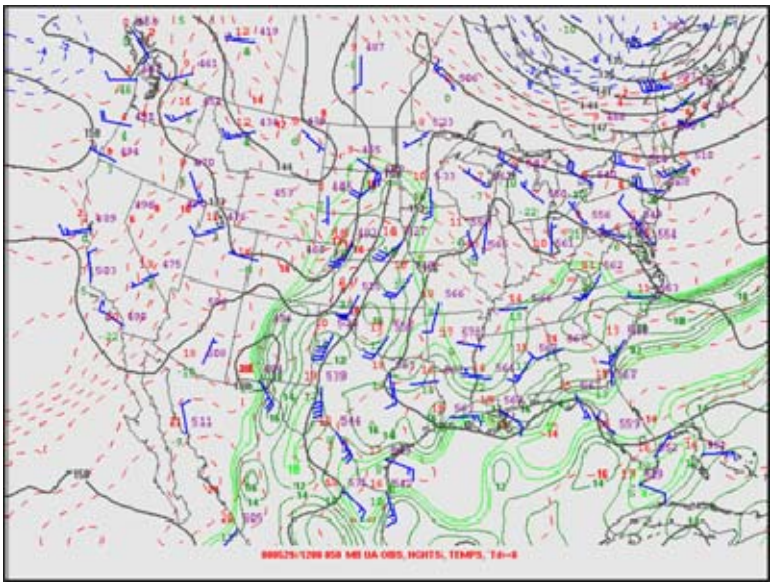


Figure 25. 080530/00Z 850 mb Analysis (From: SPC webpage, n.d.).

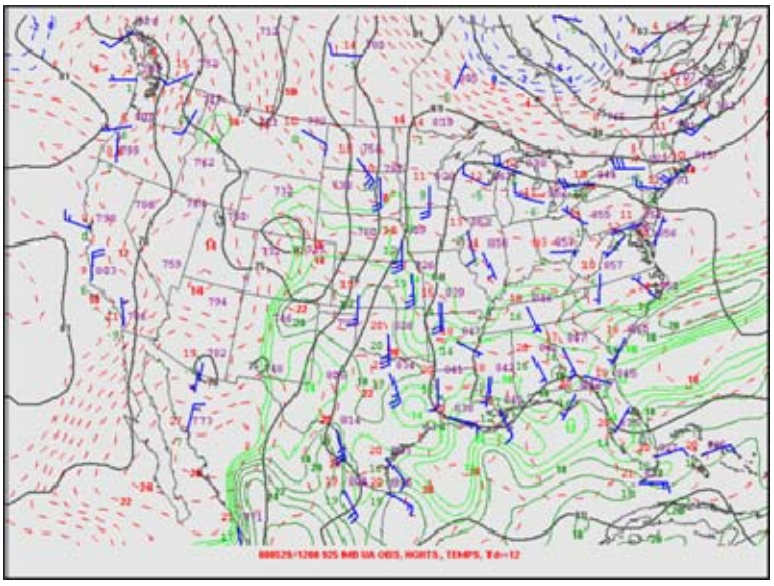


Figure 26. 080530/00Z 925 mb Analysis (From: SPC webpage, n.d.).

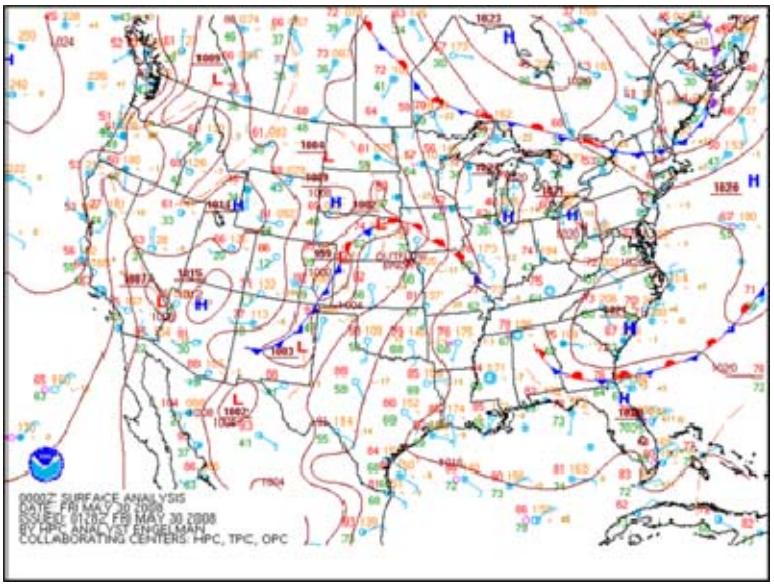


Figure 27. 080530/00Z Surface Analysis (From: SPC webpage, n.d.).

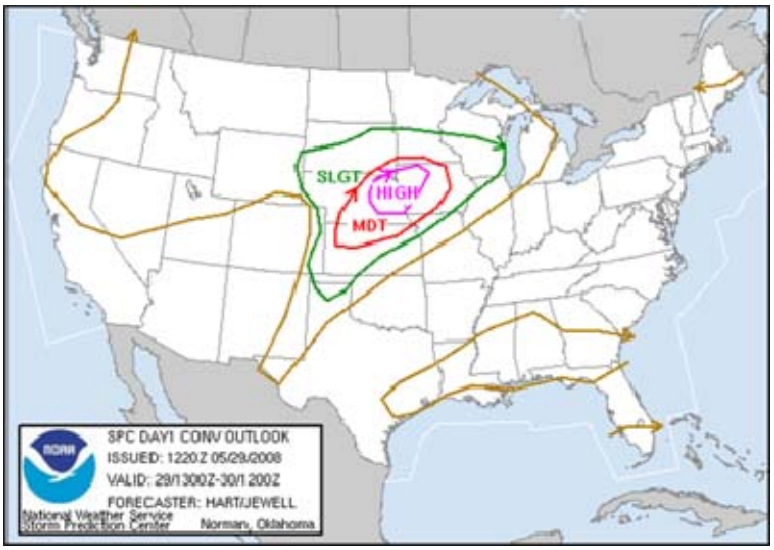


Figure 28. SPC Day 1 Convective Outlook valid May 2008 29/1300Z – 30/12Z (From: SPC webpage, n.d.).

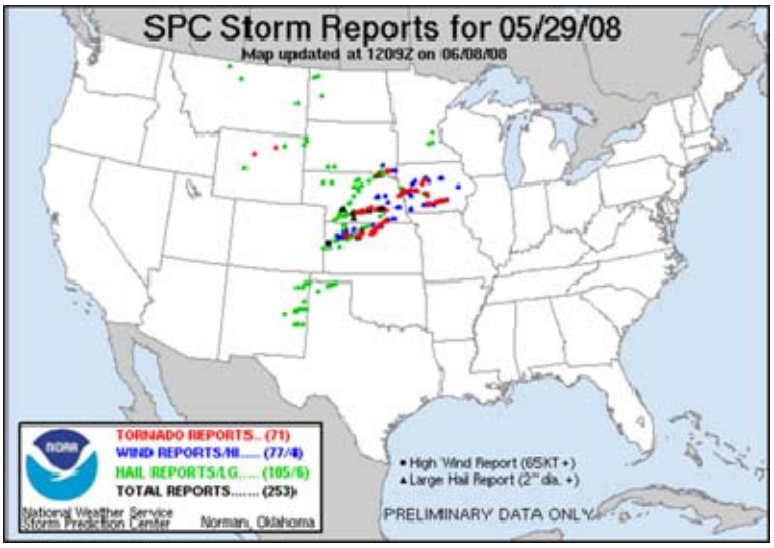


Figure 29. SPC Storm Reports Valid May 2008 29/1200Z – 30/1200Z (From: SPC webpage, n.d.).

The theta-e, LR, the respective ridges, intersections, as well as the STR, valid 30/00Z (tau = 12) are shown below in Figure 30 – Figure 35. In Figure 34, the overlap of the theta-e and LR ridges (orange and purple contours) are the forecasted ridge intersections. Figure 35 highlights these regions inside the STR as green shading. Upon applying the TELR and TELR^{+} linear discriminate analysis, the probabilistic forecasts for each are in Figure 36 – Figure 37 and the verification is plotted in Figure 38. Also evident in Figure 38, the STR appears to be a respectable delineation between verifying organized severe vs. no severe. This method attempts to add skill be increasing resolution inside this region.

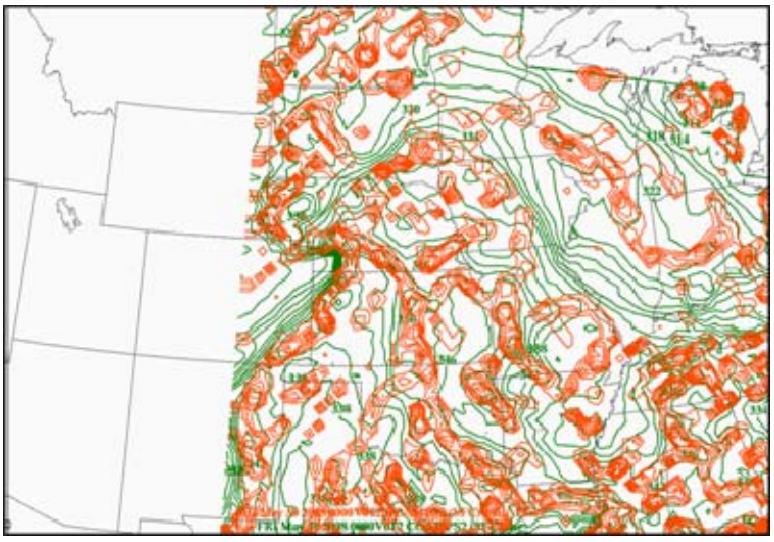


Figure 30. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Maximum Theta-e (Surface – 700 mb) (green contours) and Ridges (orange contours).

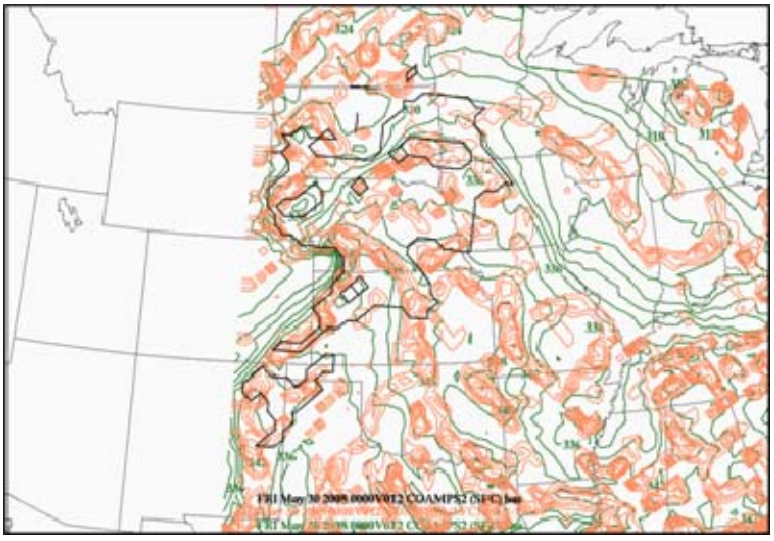


Figure 31. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Maximum Theta-e (Surface – 700 mb) (light green contours), Ridges (light orange contours) and STR (black line).

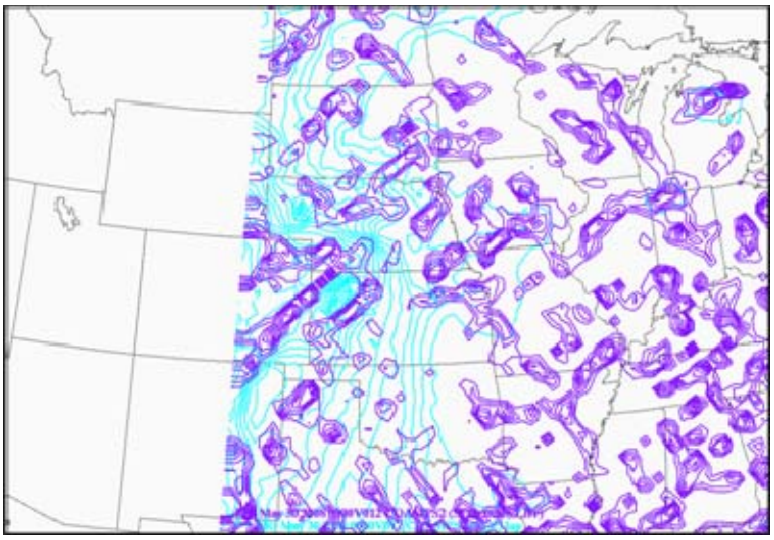


Figure 32. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Lapse Rate (500 mb to 50 mb above surface) (blue contours) and Ridges (purple contours).

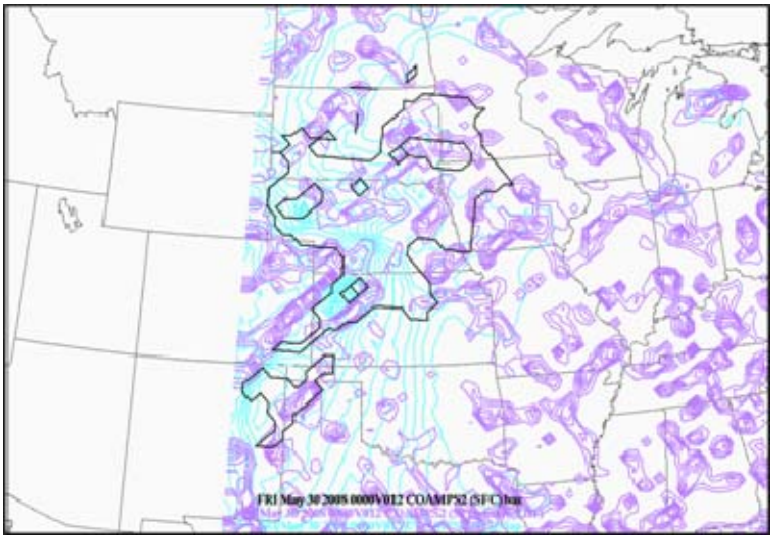


Figure 33. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Lapse Rate (500 mb to 50 mb above surface) (light blue contours), Ridges (light purple contours) and STR (black line).

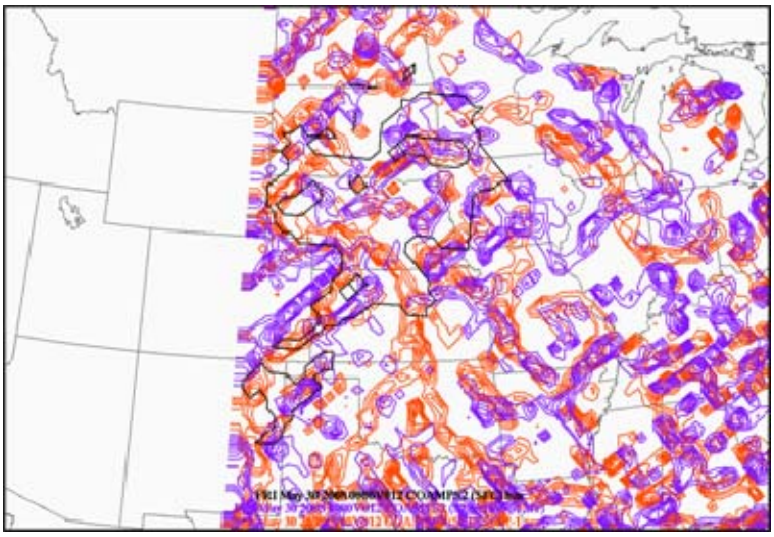


Figure 34. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). Lapse Rate (500 mb to 50 mb above surface) Ridges (purple contours), Maximum Theta-e (Surface – 700 mb) Ridges (orange contours), and STR (black line).

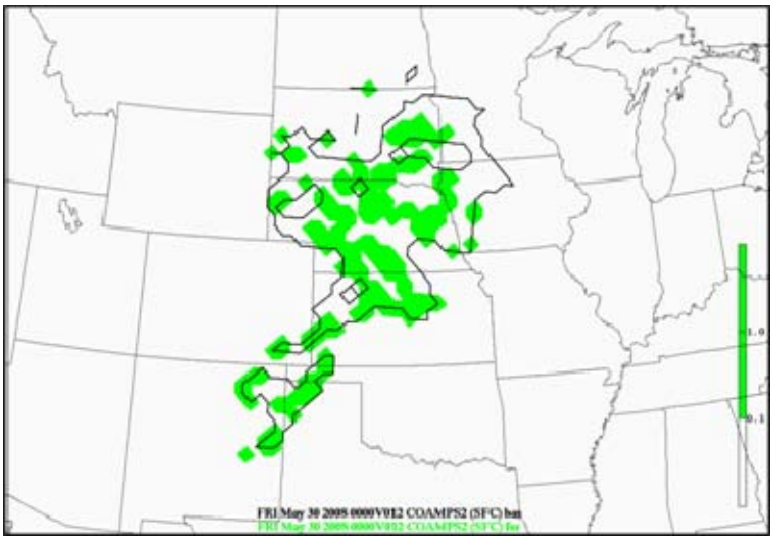


Figure 35. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z) Theta-e and LR Intersections (green area) and the STR (black line).

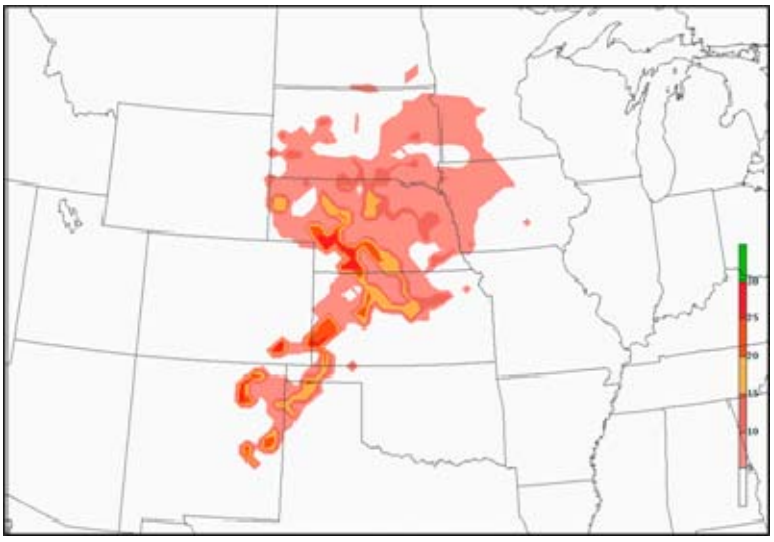


Figure 36. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). TELR Probabilistic Forecast (probabilities > 5% are shaded in 5% increments).

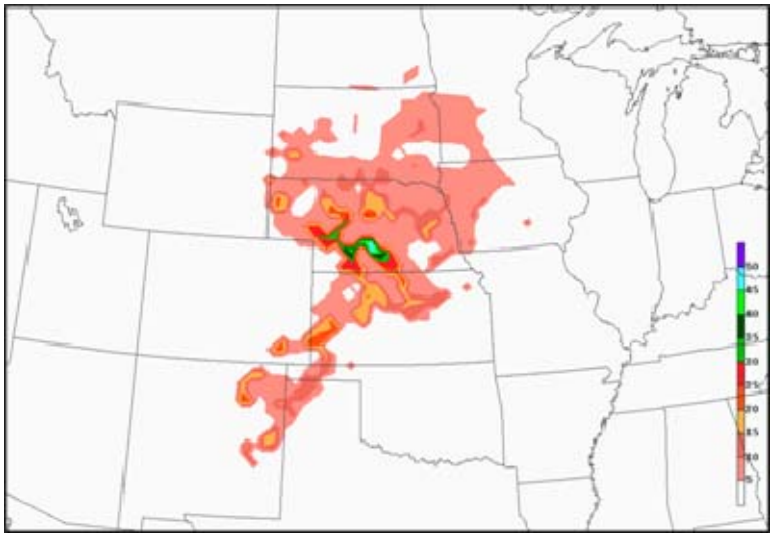


Figure 37. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). TELR^{+} Probabilistic Forecast (probabilities > 5% are shaded in 5% increments).

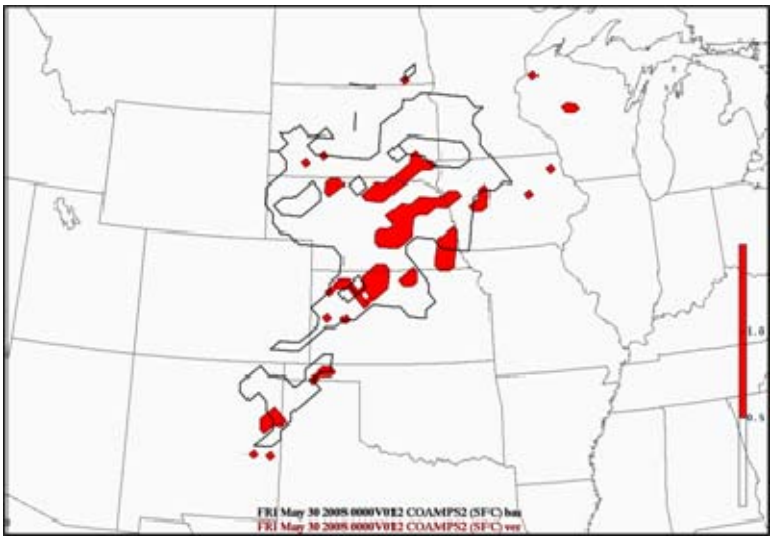


Figure 38. Verification (red area) using Radar Reflectivity Composites and Storm Reports and STR (black line) valid 080529/2345Z – 080530/0015Z.

2. TELR Forecast Statistical Results

After seeing an example of the TELR and TELR^{+} forecasts and how they were determined, the statistical results are presented to determine if the forecasts have skill.

The Brier Skill Score (BSS) for the TELR Forecast as well as the reliability, resolution and uncertainty are shown below in Figure 39 and Figure 40 (the formulas for reliability, resolution and uncertainty are in the Appendix). The two methods of calculating BSS shown in the top graph of Figure 39 and Figure 40 indicate significant positive skill according to the 95% confidence interval (CI) at all but the 24-hr forecast tau.

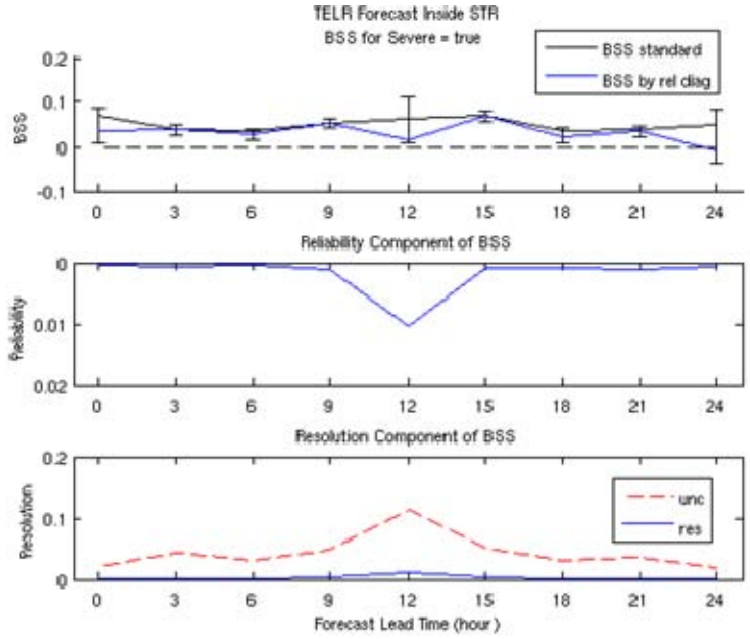


Figure 39. Brier Skill Score (BSS) for the 24-hr forecast period (top) calculated using the standard formula and $\bar{o}$ as climatology (blue line) and using reliability, resolution, and uncertainty (black line), (middle) the Reliability (rel) component of the BSS and (bottom) the Resolution (res) and Uncertainty (unc) components of the BSS for the TELR Forecast Inside the Severe Threat Region (STR).

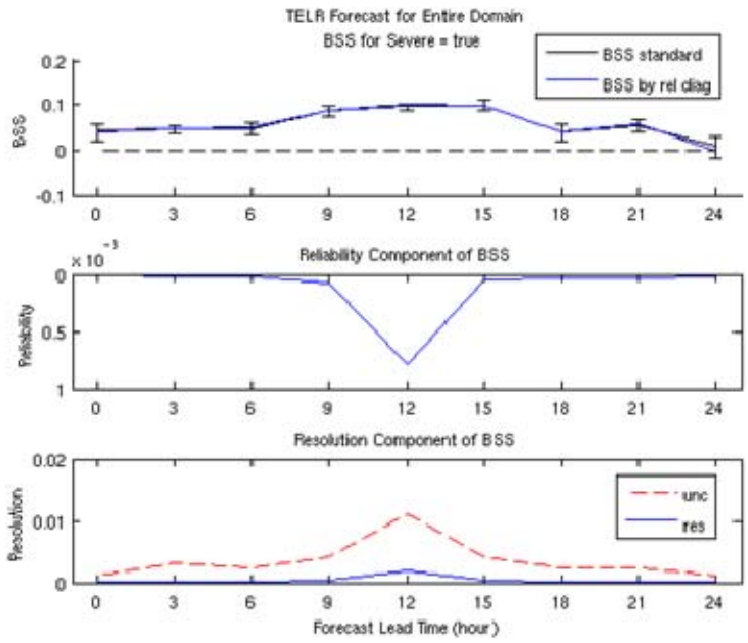


Figure 40. Brier Skill Score (BSS) for the 24-hr forecast period (top) calculated using the standard formula and $\bar{o}$ as climatology (blue line) and using reliability, resolution, and uncertainty (black line), (middle) the Reliability (rel) component of the BSS and (bottom) the Resolution (res) and Uncertainty (unc) components of the BSS for the TELR Forecast for the Entire Domain.

The analysis hour shows significant positive skill that determines validity of the forecast technique. The large 95% CI of 0.07 at the analysis hour (12Z) is a result of a relatively small sample size (4773) of severe forecasts due to the minimum in the diurnal cycle (can also be seen at tau = 24). The forecast technique indicates a positive measure of skill until the 24-hr forecast. The relatively large CI at tau = 5 (00Z) is not explained by the sample size (9601), but by the under-forecasting of severe weather at the diurnal peak (this will be expanded on later).

The BSS for the entire domain shown in Figure 40 indicates significant positive skill at the analysis hour through the 21-hr forecast as well. The skill is higher (than the STR only) at each forecast tau and the 95% CI is tighter for practically each forecast tau.

For both domains, the 12-hr forecast indicates an increase in the uncertainty term (bottom graphs in Figure 39 and Figure 40), which can drive down BSS (in the Appendix), created here by the increase in sample climatology (in the Appendix). The

sample climatology increase from 2% to 13% (shown in Figure 41) from tau = 0 to 12 corresponds to an increase in uncertainty from 0.02 to 0.11 (shown graphically in Figure 42 below).

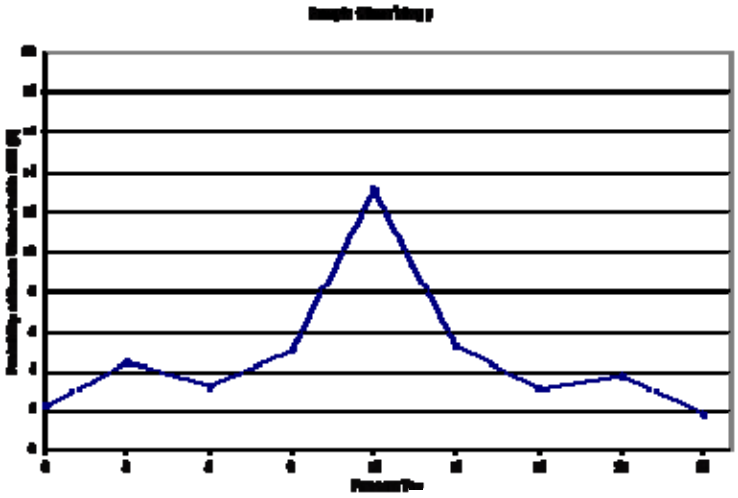


Figure 41. Sample Climatology for the 24-hr Forecast Period of Severe Verification Inside the STR.

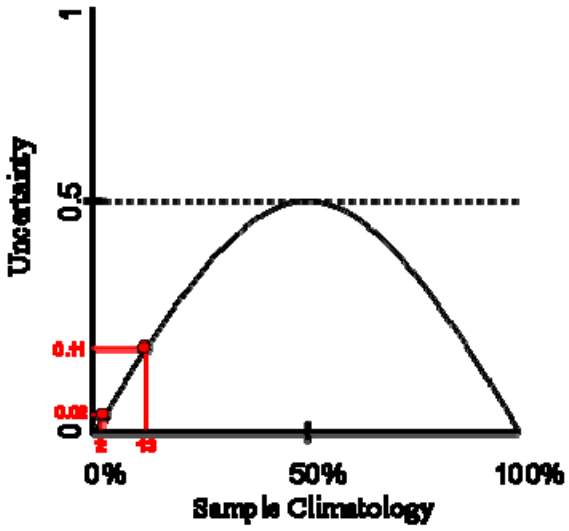


Figure 42. Uncertainty vs. Sample Climatology.

The reliability diagram is a useful tool to graphically comprehend the full joint distribution of forecasts and observations at a particular forecast tau (Wilks, 1995). The reliability diagram can also be use to diagnose calibration needs and strengths and weaknesses of the forecast. The reliability diagram (left) as well as the probability bin usage (right) valid at tau = 12 for the TELR Forecast is shown below in Figure 43 and Figure 44.

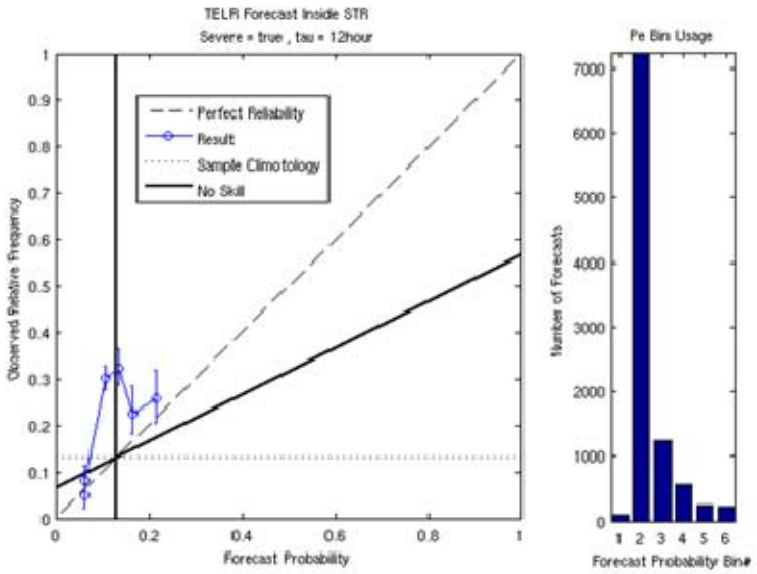


Figure 43. Reliability Diagram (left) and Forecast Probability Bin Usage (right) (tau = 12) for the TELR Forecast Inside the Severe Threat Region (STR).

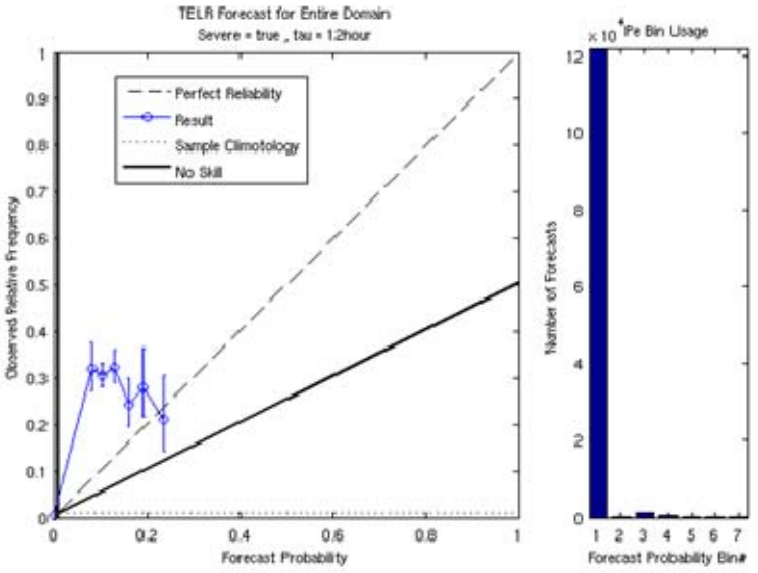


Figure 44. Reliability Diagram (left) and Forecast Probability Bin Usage (right) (tau = 12) for the TELR Forecast for the Entire Domain.

The first thing to note is the large number of forecasts in the second bin of Figure 43 (right plot). Of these 7242 forecasts, 6,878 are non-intersections ('no' forecasts) and 364 are intersections ('yes' forecasts) with a forecast probability between 0.0631 (6.31%) and 0.093. The overwhelming number of non-intersections inside the STR with an assigned forecast probability of 0.0631 (at tau = 12) will contribute to an increase in the BSS. The same effect occurs when analyzing the entire domain (Figure 44) except the majority of forecasts are the gridpoints outside the STR (approx 120,000) that have an assigned forecast probability of 0 and an observed relative frequency (ORF) of nearly zero.

There was also a tendency for the technique to under forecast the probability of severe at tau = 12 for all bins except bin six. This could be explained by the diurnal maximum in severe weather occurrence at tau = 12 (00Z). Since the linear discriminate analysis trained on the analysis and all forecast taus (00 – 24-hr), the independent dataset forecast will under-forecast severe at max heating (evident in Figure 43 and Figure 44) and will over-forecast severe for morning forecast taus. The over-forecasting of severe is evident in the reliability diagrams in Figure 45 and Figure 46 for the tau = 6 and 24 (18Z

and 12Z). It is evident by the under and over forecasting of severe at different taus that the forecast hour (time of day) would be a beneficial predictor to add to the discriminate analysis.

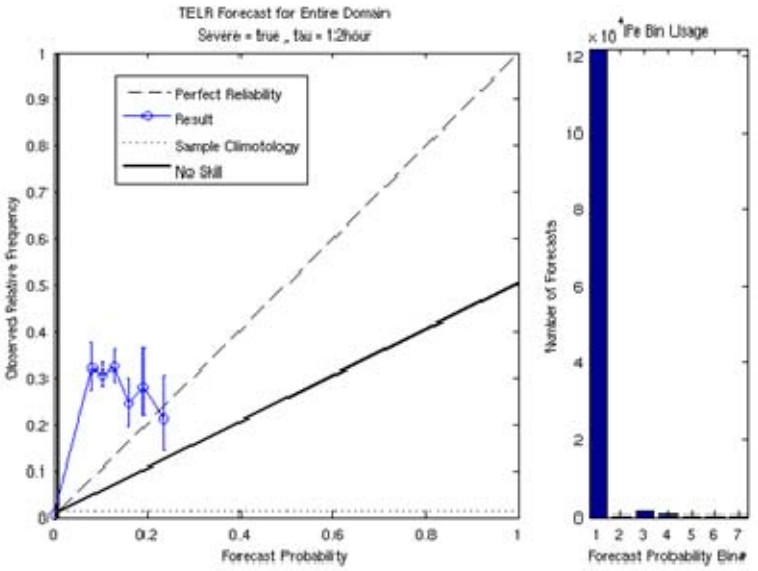


Figure 45. Reliability Diagram (left) and Forecast Probability Bin Usage (right) (tau = 06) for the TELR Forecast Inside the Severe Threat Region (STR).

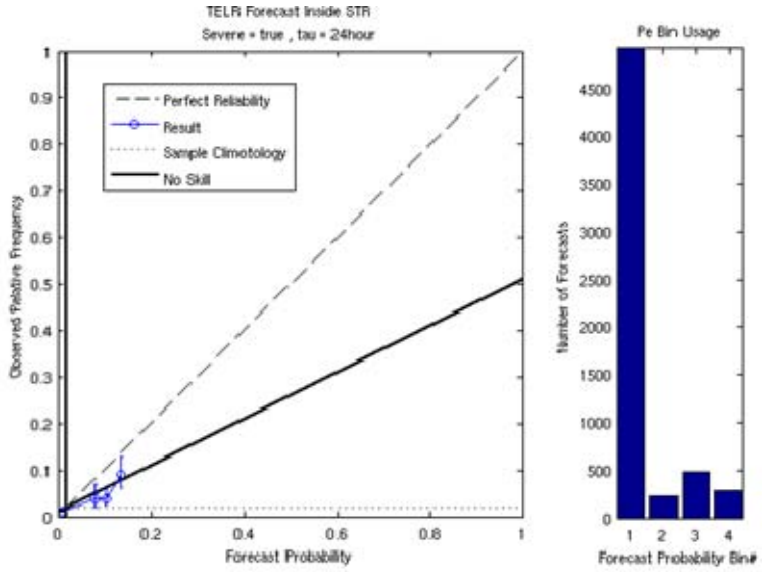


Figure 46. Reliability Diagram (left) and Forecast Probability Bin Usage (right) (tau = 24) for the TELR Forecast Inside the Severe Threat Region (STR).

The Relative Operating Characteristic (ROC) is another graphical display of a discrimination-based forecast (severe or no severe), but has different attributes than the BSS. The idea behind the ROC diagram is the hypothetical decision of taking action or not based on a probabilistic forecast of a dichotomous event. By varying the probabilistic threshold at which the decision to act is based, a new 2 × 2 contingency table is created (as in Figure 21) where the 'Forecast' (Y or N) now becomes Act (Y or N). The false alarm rate, $\frac{b}{b+d}$, and the hit rate, $\frac{a}{a+c}$, at the different thresholds are now the abscissa and ordinate of the graphical representation. The ROC diagrams for the TELR forecast for both domains are shown below in Figure 47 and Figure 48. The graphical display of positive skill on a ROC diagram is indicated by the integrating the area below the traced line (also shown as A in the Appendix) and comparing with the area of the random forecast. The random forecast traces the line along the 45° diagonal connecting the points (0, 0) and (1, 1) (Wilks, 1995). The integrated area beneath this line is 1 and represents no skill. The better the forecast becomes, the closer it approaches the upper left corner (false alarm rate = 0% and hit rate = 100%).

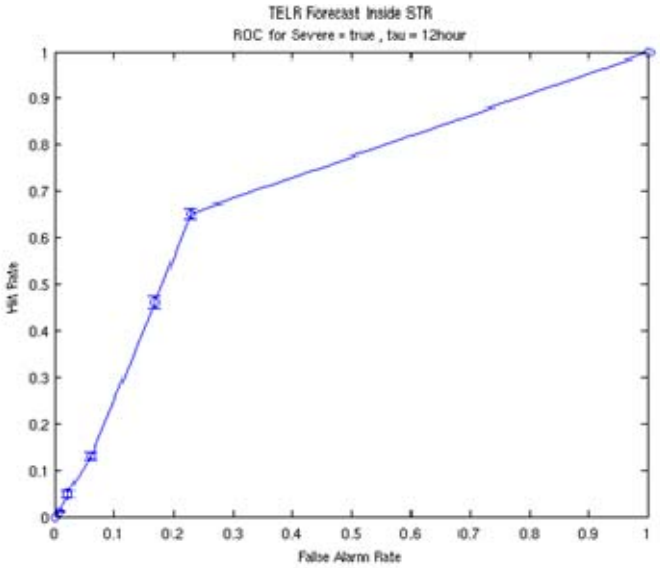


Figure 47. ROC Diagram (tau = 12) for the TELR Forecast Inside the Severe Threat Region (STR).

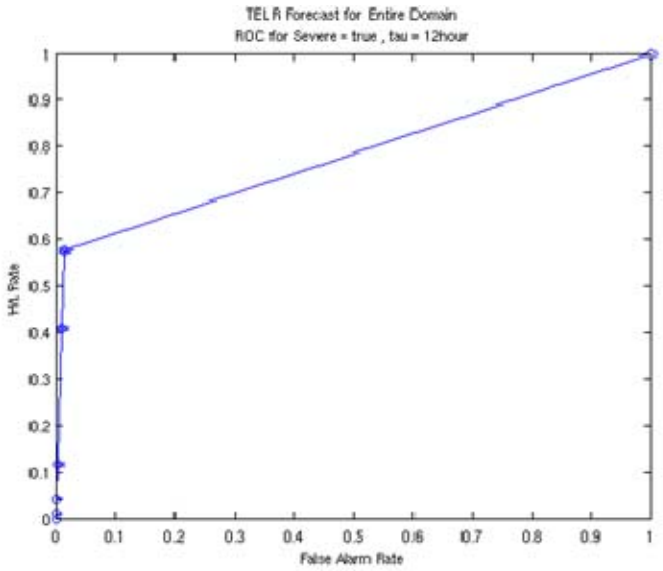


Figure 48. ROC Diagram (tau = 12) for the TELR Forecast for the Entire Domain.

The ROC diagram inside the STR indicates the ability to discriminate between severe and non-severe events as the traced line is extending toward the upper-left corner. As we increase the domain (Figure 48) the false alarm rate for all probability thresholds decreases, indicating the rarity of the forecast event on this domain. The maximum hit rate also dropped from approximately 65% in the STR to 58% across the entire domain.

The ROC skill score (ROCSS) (formula found in the Appendix) is a measure of forecast skill with respect to a random forecast. By the ROC diagrams at tau = 12, there was a positive indication of skill. Figure 49 and Figure 50 display the ROCSS for the 24-hr forecast period for both domains.

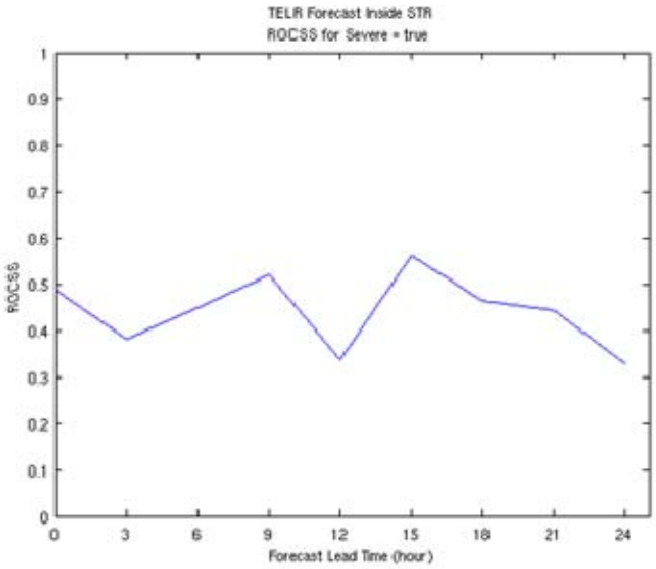


Figure 49. Relative Operating Characteristic Skill Score (ROCSS) for the 24-hr forecast period for the TELR Forecast Inside the Severe Threat Region (STR).

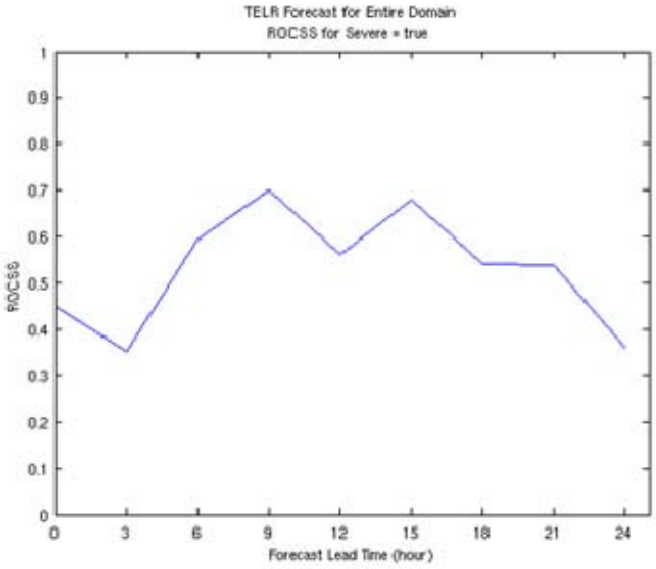


Figure 50. Relative Operating Characteristic Skill Score (ROCSS) for the 24-hr forecast period for the TELR Forecast for the Entire Domain.

The ROCSS inside the STR indicates positive skill throughout the entire forecast, but the ROCSS over the entire domain appears to display greater skill at the highest climatologically favorable severe weather hours (Figure 50). This indicates the forecast method resolves the event the best when there is more severe convection present.

The value score (VS) is a measure of forecast utility for different decision makers (Wilks, 1995). The decision to take action based on a probabilistic forecast can be determined by the risk tolerance of the user. This measurement, like the ROC diagram, is based on signal detection theory. The VS is a measure of potential economic value as a function of the risk tolerance of the user (or cost-to-loss ratio). The VS for the TELR forecast for each domain is shown below in Figure 51 and Figure 52.

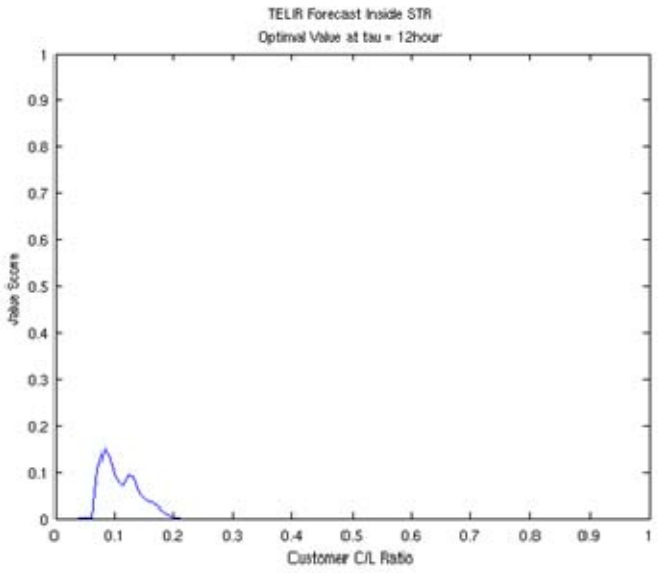


Figure 51. Value Score (VS) (tau = 12) for the TELR Forecast Inside the Severe Threat Region (STR).

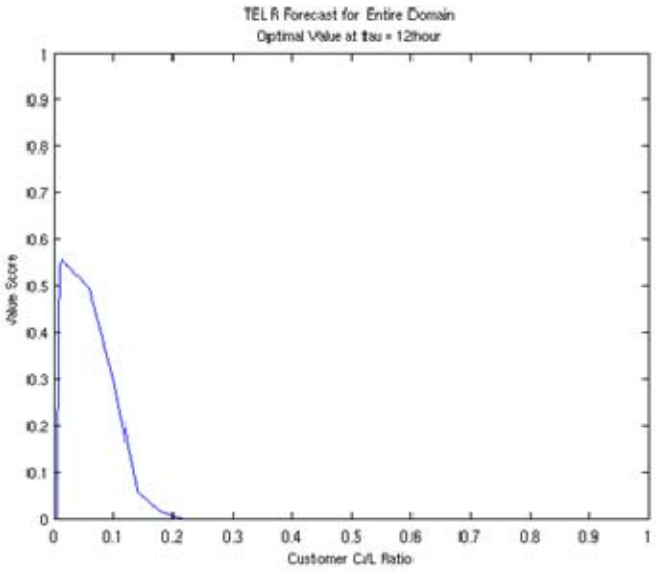


Figure 52. Value Score (VS) (tau = 12) for the TELR Forecast for the Entire Domain.

The TELR forecast inside the STR (Figure 51) indicates a potential increase in value over the sample climatology inside the STR for customers with a risk tolerance (or cost-to-loss ratio) between 0.07 – 0.2. This indicates risk-adverse customers will slightly benefit from the TELR forecast inside the STR. If the TELR forecast is applied over the entire domain, the value increases dramatically for a larger range of risk-adverse customers. The addition of value to only risk adverse users is expected. A perfect severe forecast would only benefit customers that are risk adverse (0 – 0.3 cost-to-loss ratio) because the low probability of the event.

3. TELR$^+$ Forecast Statistical Results

In an attempt to increase skill in the forecast method, MUCAPE and 0 – 6 km shear were added as predictors (in addition to theta-e, LR, and ridge magnitudes) comprising the linear discriminate analysis to make the TELR$^+$ forecast. The TELR$^+$ forecast is comprised of six predictors where the TELR used only four. The statistical analysis was re-run and the same skill metrics were created as in the TELR forecast. The first indicator of skill to examine is the BSS in Figure 53 and Figure 54.

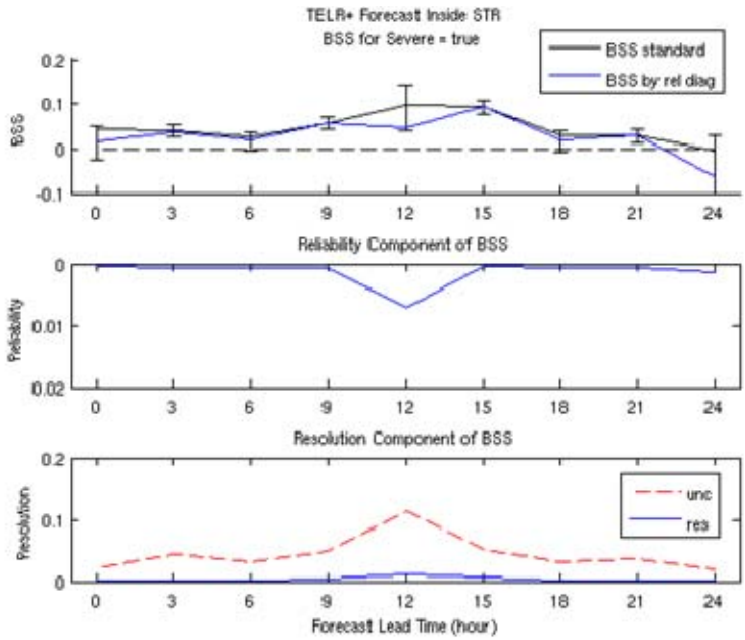


Figure 53. Brier Skill Score (BSS) for the 24-hr forecast period (top) calculated using the standard formula and $\bar{o}$ as climatology (blue line) and using reliability, resolution, and uncertainty (black line), (middle) the Reliability (rel) component of the BSS and (bottom) the Resolution (res) and Uncertainty (unc) components of the BSS for the TELR^{+} Forecast Inside the Severe Threat Region (STR).

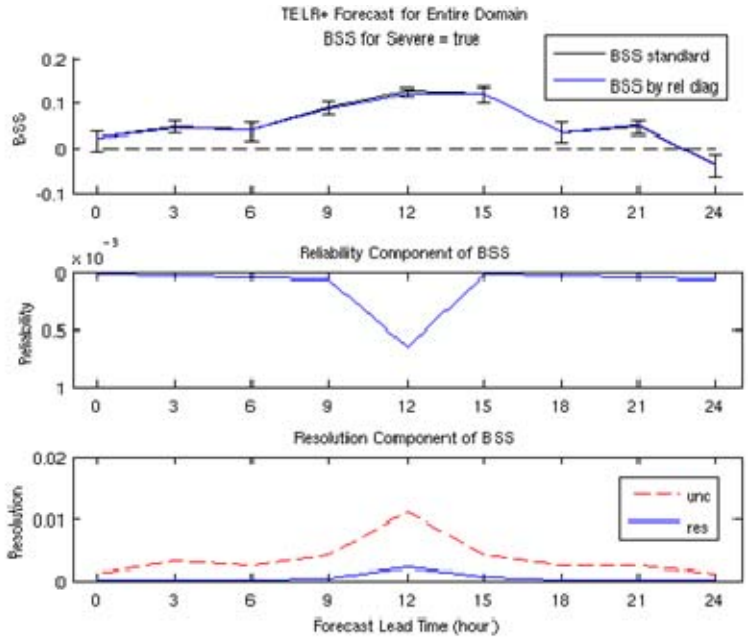


Figure 54. Brier Skill Score (BSS) for the 24-hr forecast period (top) calculated using the standard formula and $\bar{o}$ as climatology (blue line) and using reliability, resolution, and uncertainty (black line), (middle) the Reliability (rel) component of the BSS and (bottom) the Resolution (res) and Uncertainty (unc) components of the BSS for the TELR^{+} Forecast for the Entire Domain.

The TELR^{+} forecast at the analysis hour (indicated by 95% CI) shows no significant positive skill. An interesting feature to note is the increase in skill during max heating (tau = 9 – 15). The addition of MUCAPE as a predictor in the TELR^{+} forecast could explain the diurnal increase in skill at max heating. Corresponding to the peak in sensible heat flux at the surface, the most-unstable parcels tend to become surface-based and show an increase in MUCAPE. This diurnal increase in MUCAPE and severe weather could be the main reason for the increase in skill of the TELR^{+} forecast in the late afternoon.

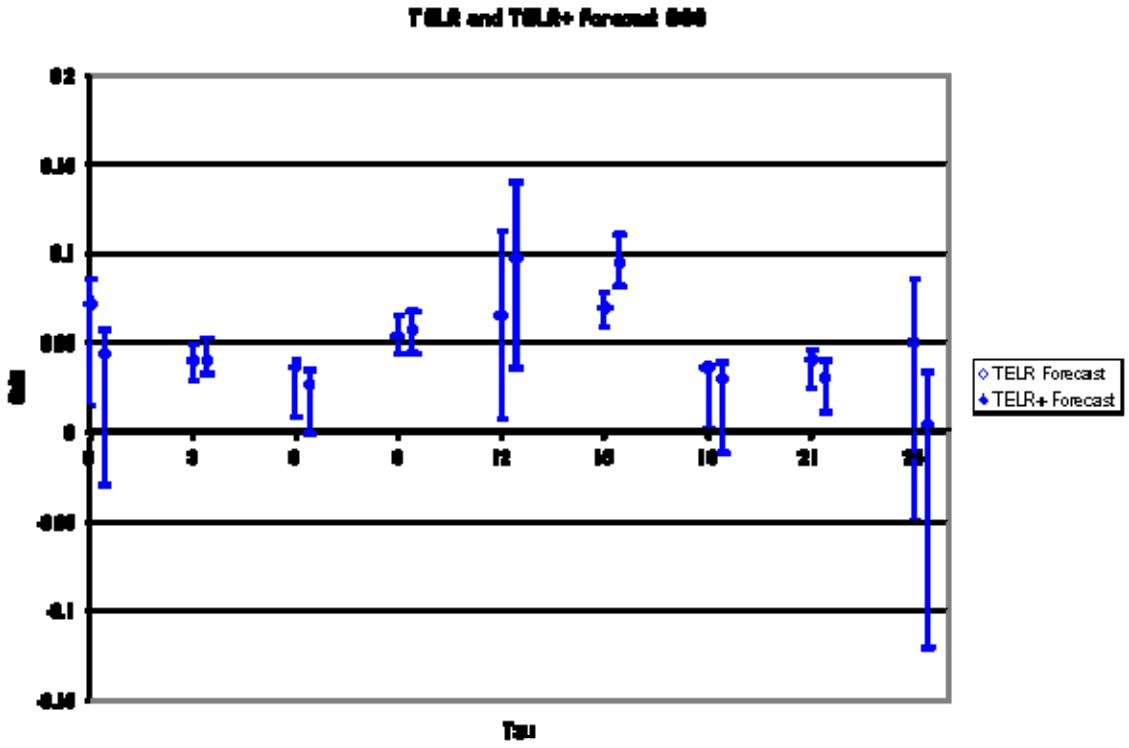


Figure 55. The BSS for the 24-hr Forecast Period TELR (blue circles) and TELR^{+} (filled blue circles) Forecasts with 95% Confidence Interval.

As evident in Figure 55, there is no statistical difference between the TELR and TELR^{+} forecasts (except at tau = 15), although there are a couple interesting features to note. The first is that TELR^{+} has lower skill (not significant) than TELR at the analysis hour and greater skill in the late afternoon (only significant at tau = 15), hinting again at a stronger correlation to the diurnal insolation trend than the TELR forecast. As in the TELR forecast, this should indicate a over-forecasting of severe weather at tau = 12 and an under-forecasting of severe at tau = 00. The reliability diagrams for each are below in Figure 56 and Figure 57, indicating under and over-forecasting, respectively.

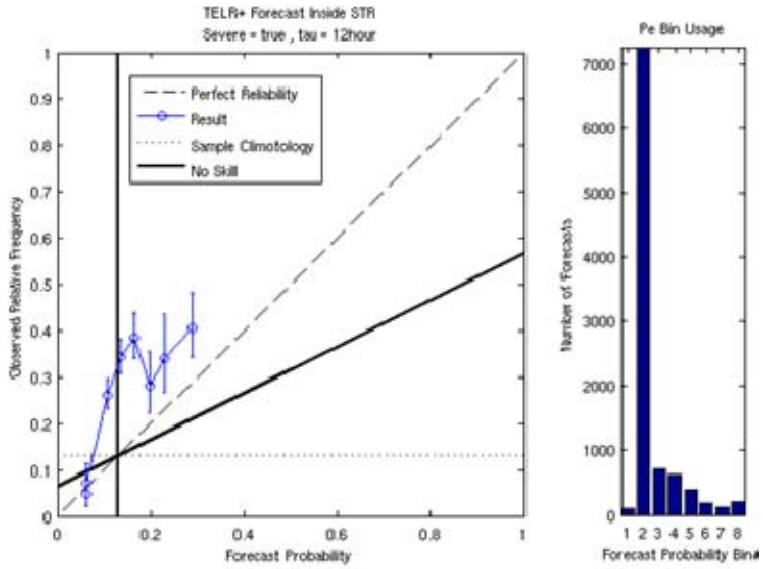


Figure 56. Reliability Diagram (left) and Forecast Probability Bin Usage (right) (tau = 12) for the TELR^{+} Forecast Inside Severe Threat Region (STR).

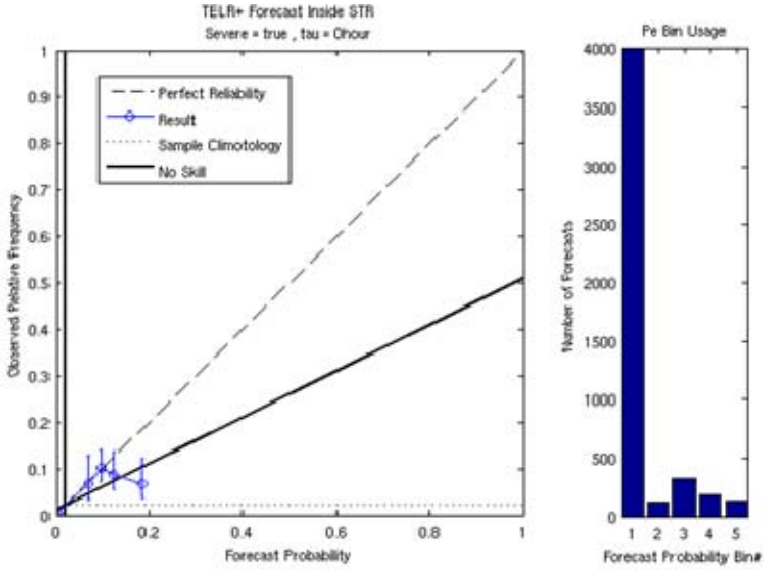


Figure 57. Reliability Diagram (left) and Forecast Probability Bin Usage (right) (tau = 00) for the TELR^{+} Forecast Inside Severe Threat Region (STR).

The other difference noted between the TELR and TELR^{+} metrics was the VS. The VS for TELR^{+} forecast over both domains is shown in Figure 58 and Figure 59.

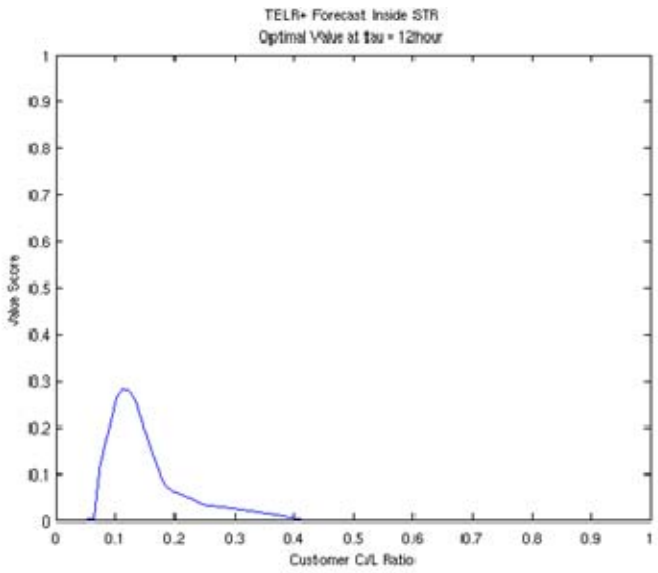


Figure 58. Value Score (VS) (tau = 12) for the TELR^{+} Forecast Inside the Severe Threat Region (STR).

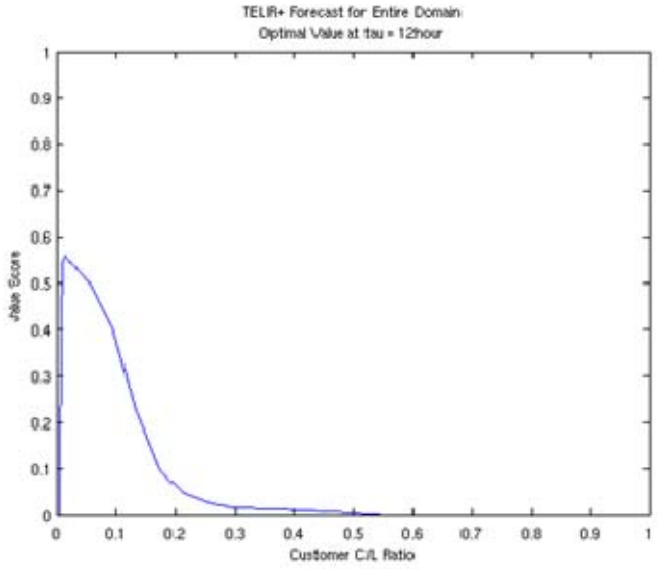


Figure 59. Value Score (VS) (tau = 12) for the TELR^{+} Forecast for the Entire Domain.

The VS for the TELR^{+} forecast inside the STR indicates a larger potential economic value for a larger range of customers (cost-to-loss ratio of 0.08 – 0.5).

Increasing the domain to the entire grid expands the potential value to extremely risk-adverse customers (as does the TELR forecast), but still provides potential value to the less risk-adverse customer.

4. Discussion of Results

a. Experiment Design Limitations

There are many key understandings to be pointed out. First, the method of experiment design could be flawed. Using model forecasts to verify the application of a new forecast technique is premature. This assumes a sound forecast technique that is repeatable. The forecast event in question is organized severe weather that has a life cycle of as little as 1 – 2 hr. Since the limit of predictability for mesoscale features is on the temporal timescale of the event lifespan, what is a reasonable limit of predictability of these mesoscale features? Mesoscale predictability is inherently less than the synoptic-scale due to the three-dimensional mixing vs. two-dimensional in the synoptic-scale (Ray 1986). The verification method attempted to account for model error growth by verifying severe weather within a radius of approximately 54 km from the forecast gridpoint.

In order to properly determine if the method works, the rawinsonde observation (RAOB) network analysis combined with surface and profiler observations would give an approximation to the true state of the theta-e and LR fields. The mesoscale structure portrayed in the 12-km operational NWP would not necessarily be available through this analysis, but would that be detrimental? This leads to the question, how accurately must we observe the atmosphere initially to have the fine scale structure produced by the 12-km operational NWP model? Are the high resolution features in the theta-e and LR parameters artificial because we currently cannot observe the atmosphere on the 12 km scale? This will be discussed further in this chapter. Also, the analysis of the hourly Rapid Update Cycle (RUC) model can be used as an approximation to the truth to test the validity of the technique.

b. Uncertainty Portrayal

The application of any skilled forecast method using an individual NWP forecast member has inherent flaws. It is determined that the forecast technique has skill, but using just one NWP solution with erred initial and boundary conditions (IC and BC) and flawed model physics could lead to model error growth in the forecast parameters that render the method useless. In order to portray the uncertainty in the severe forecast, a multitude of solutions with varied ICs, BCs, and model physics are required to attempt to capture the entire spectrum of solutions. Each solution would be used to train the discriminate analysis to produce a robust probabilistic severe forecast based on estimated uncertainty.

c. Wavelength of Severe Signal

When the idea of this forecast technique was born, hand analysis of the intersections of the 850 mb theta-e ridge and 500 mb – 925 mb temperature differential were analyzed. The human forecaster uses pattern recognition as well as meteorological situational awareness to weight important features (intersections) as well as particular wavelength features heavier than others. Automating this process is not an easy task. For example, the small scale perturbations in the theta-e ridge are weighted just as heavily as the main LLJ axis. Could this be a flaw in the design of the experiment? The ridge intersections inside the STR for the example day (Figure 35) show mesoscale structure, but do we have predictability at this small spatial scale? Also, how realistic is the mesoscale structure evident in the model fields? Currently in the CONUS, the average distance between RAOB sites is approximately 400 km and the average distance between surface observations is roughly 50 km. These observations assimilated into the initial guess of the NWP model resolves wavelength features in the upper-air of approximately 2000 km and at the surface approximately 250 km. There are other sources of data assimilated into NWP, but can they accurately resolve (analyze) the mesoscale structure evident in the 12-km operational NAM-WRF analysis? What about after the 12-km NAM-WRF is interpolated to the 32-km grid? Can the meso-β structure in the forecast parameters be assimilated? What if the smaller scale perturbations in the

theta-e and LR fields were filtered resulting in the quasi-geostrophic signal? Will the ridge intersections look more like the hand analysis intersections? Will there be any difference in skill by looking at the meso-β (20 – 200 km) to meso-α (200 – 2000 km) scale? By examining the dominate background structure, will the predictability of this method be extended?

To test this, a basic nine-point averaging technique will be used to filter the small scale features. The procedure averages each gridpoint with the surrounding eight gridpoints, all weighted equally. A 1-D example of the averaging method is displayed in Figure 60 below. The 'No Filter' (red line) indicates the minimum resolvable wavelength on the 32-km model grid. As the averaging technique (filter) is applied multiple times (green and blue lines), the amplitude of the waves at this particular wavelength (160 km) are essentially eliminated, leaving primarily the meso-α wavelength signal.

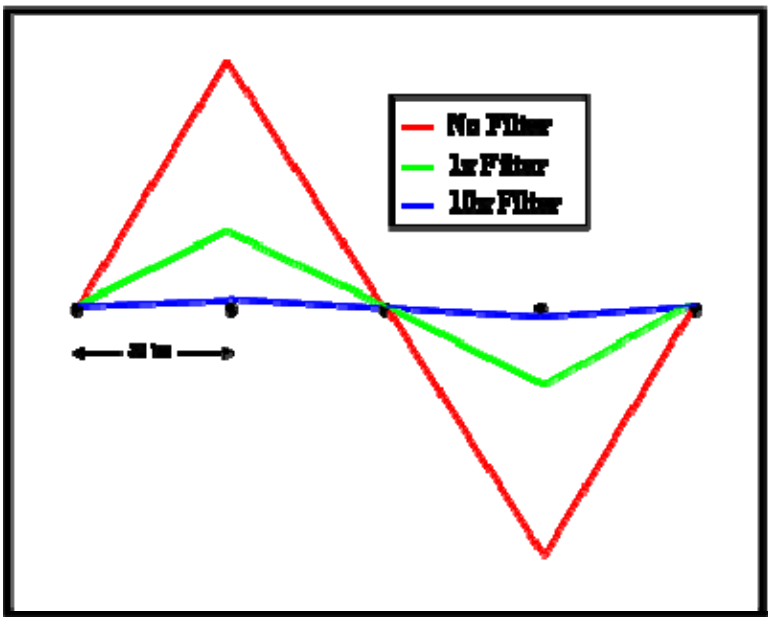
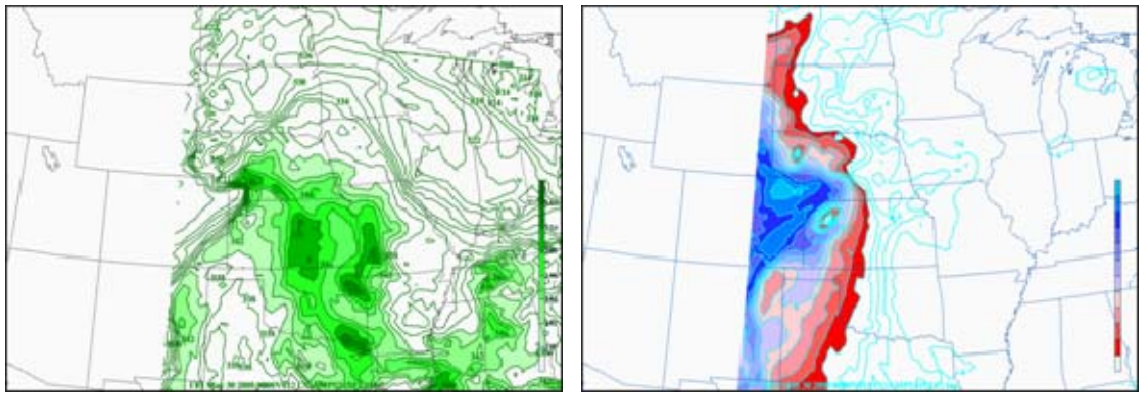


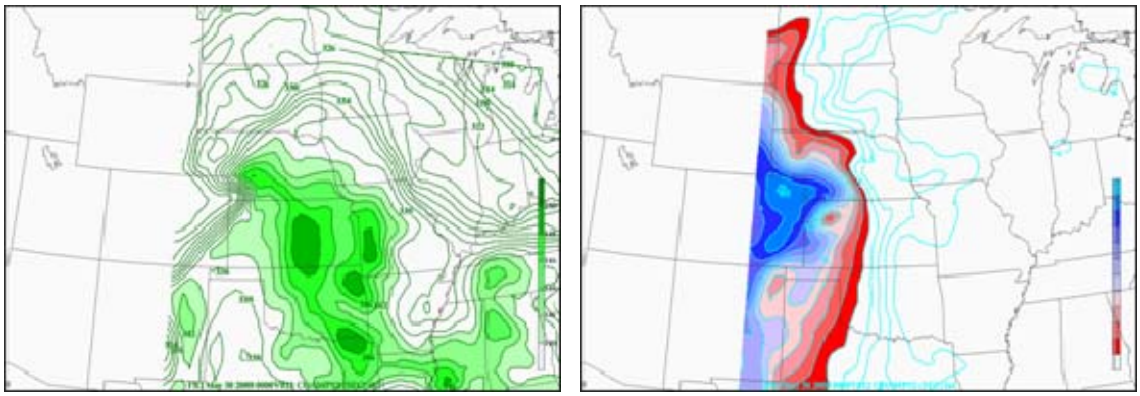
Figure 60. Example of 1-D Averaging Technique.

Each time the averaging method is applied, the larger wavelength features become dominate as shown by the theta-e and LRs shown in Figure 61 – Figure 64.



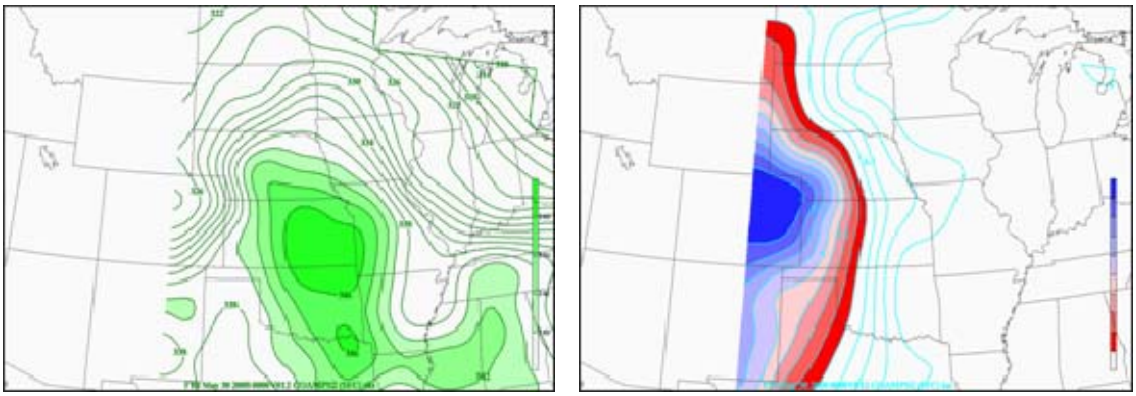
(a) (b)

Figure 61. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). No Filter – (a) Theta-e (green contours with > 340 K highlighted) and (b) LR (blue contours with > 7 deg C/km highlighted).



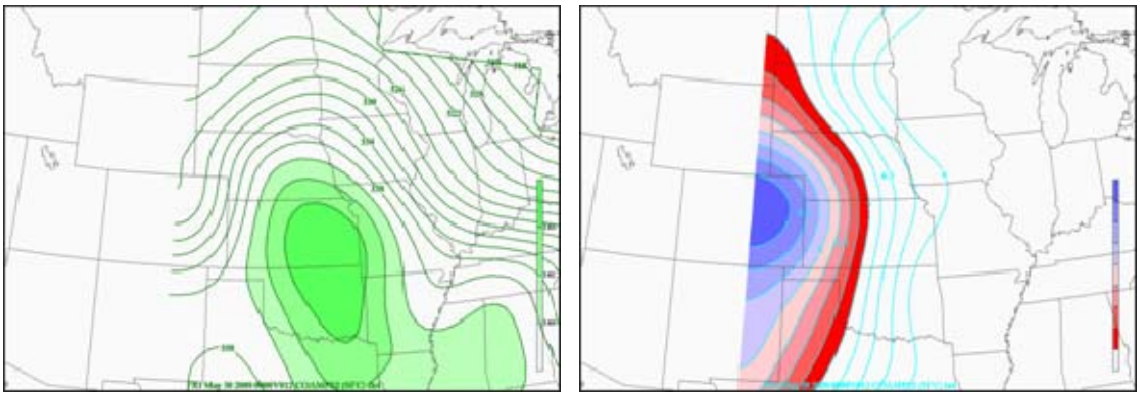
(a) (b)

Figure 62. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 1× Filter – (a) Theta-e (green contours with > 340 K highlighted) and (b) LR (blue contours with > 7 deg C/km highlighted).



(a) (b)

Figure 63. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter – (a) Theta-e (green contours with > 340 K highlighted) and (b) LR (blue contours with > 7 deg C/km highlighted).



(a) (b)

Figure 64. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 40× Filter – (a) Theta-e (green contours with > 340 K highlighted) and (b) LR (blue contours with > 7 deg C/km highlighted).

The averaging routine was applied multiple times and the results input into the dependent and independent forecast process described in Chapter III. The intersections for each forecast are shown below in Figure 65 – Figure 68. The more

times the averaging technique is applied, the more the mesoscale ridge-crossing structure conglomerates into a broad area-wide theta-e and LR ridge crossing.

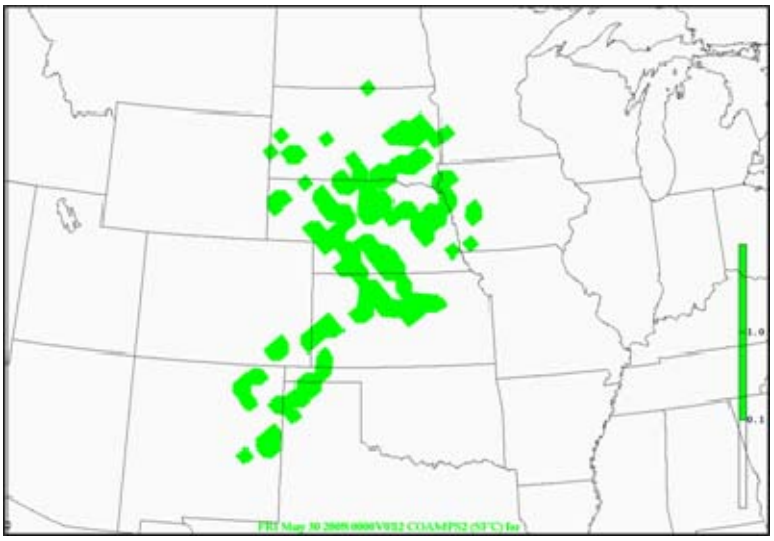


Figure 65. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). No Filter – Theta-e and LR Intersections (green area).

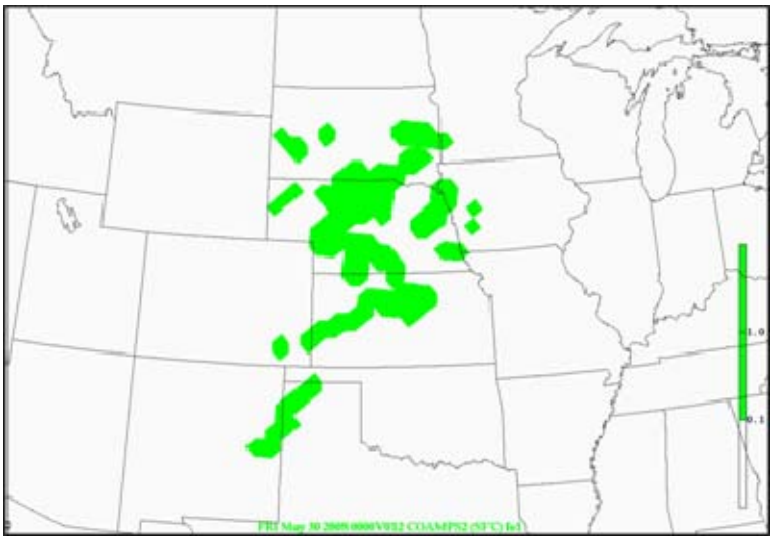


Figure 66. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 1× Filter – Theta-e and LR Intersections (green area).

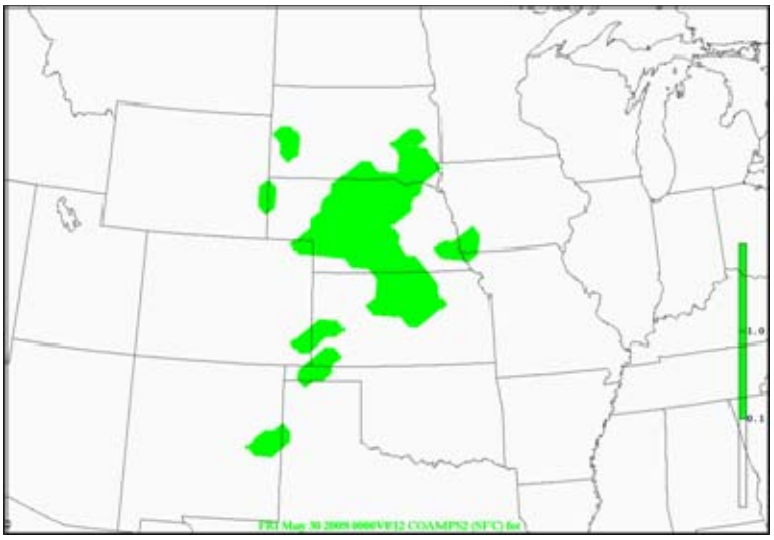


Figure 67. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter – Theta-e and LR Intersections (green area).

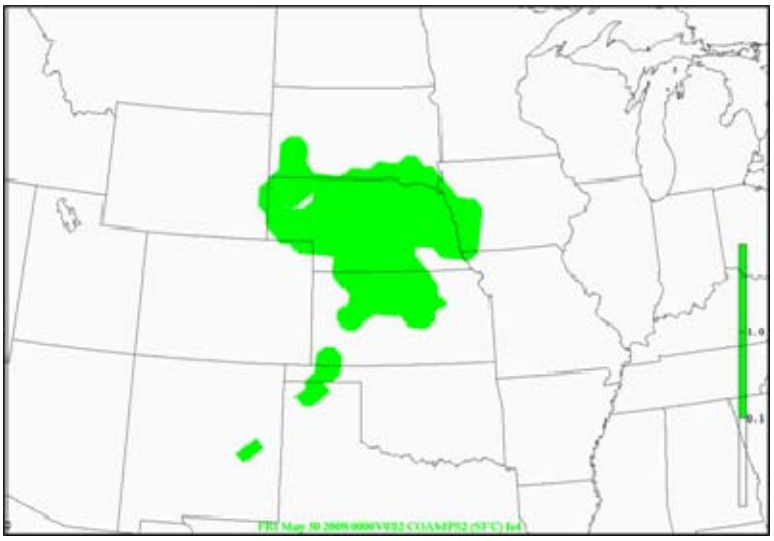


Figure 68. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 40× Filter – Theta-e and LR Intersections (green area).

The 10× filter forecast (Figure 67) appears to look like typical regions highlighted for severe weather such as severe thunderstorm or tornado watch boxes issued by the SPC (see Figure 69 for 10× filter forecast and severe and tornado watches valid at 00Z). The typical watch issued by the SPC covers about 25,000 square miles, or

half the size of Iowa (SPC webpage, n.d.). All fields used in creating and verifying the 10× filtered forecast are shown below in Figure 70 – Figure 74.

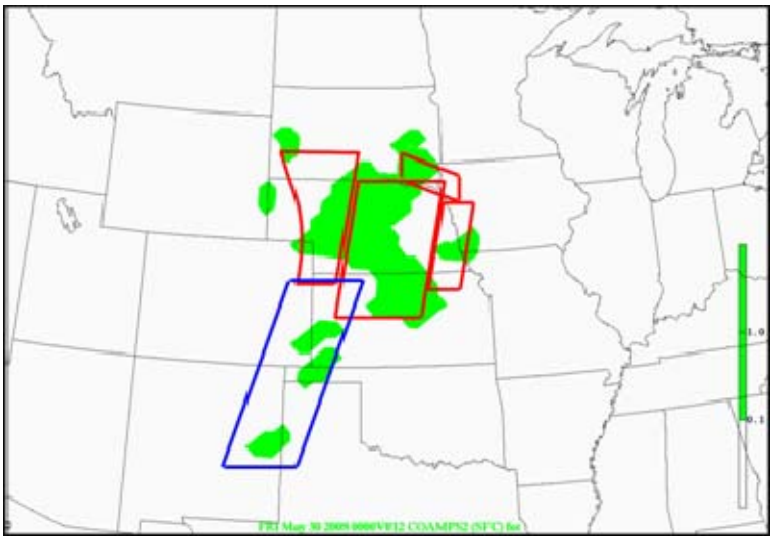
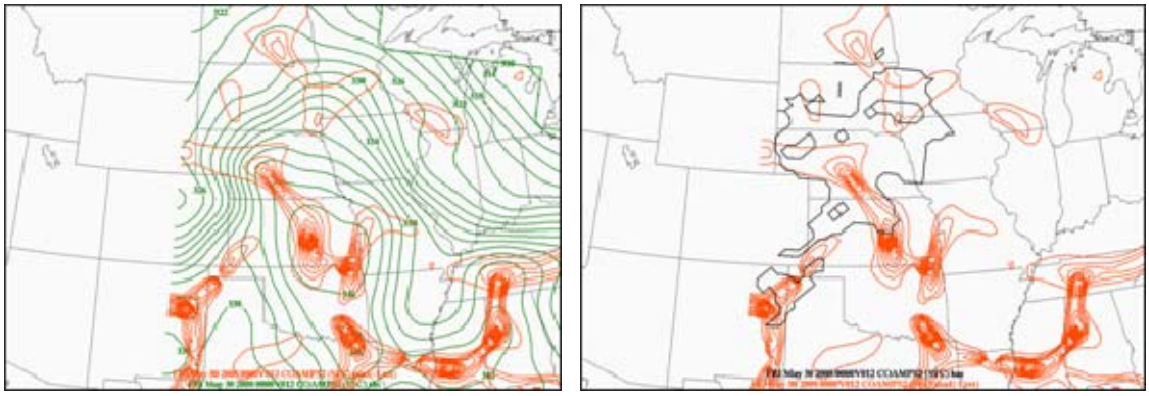
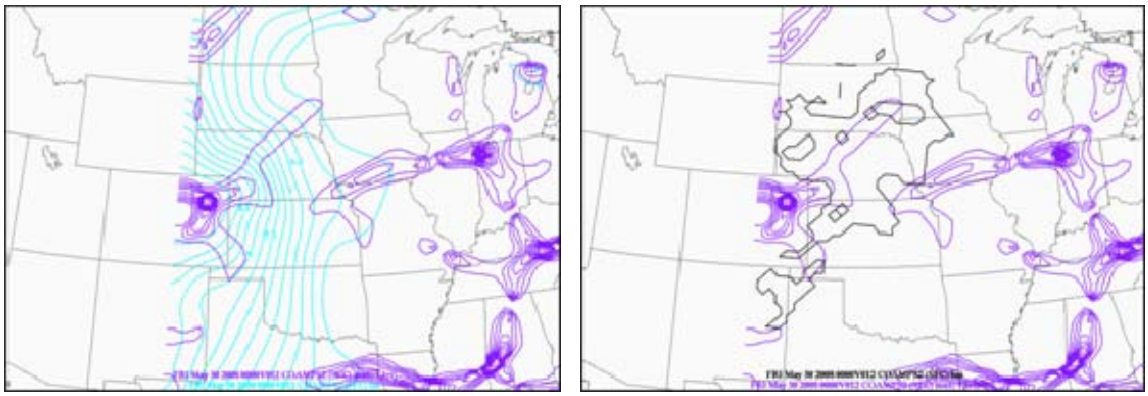


Figure 69. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter – Theta-e and LR Intersections (green area) and Tornado (red) and Severe (blue) Watch Boxes Issued by the SPC valid at 30/00Z.



(a) (b)

Figure 70. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter – (a) Theta-e (green contours) and Ridges (orange contours) and (b) Theta-e Ridges and STR (black line).



(a) (b)

Figure 71. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter – (a) LR (blue contours) and Ridges (purple contours) and (b) LR Ridges and STR (black line).

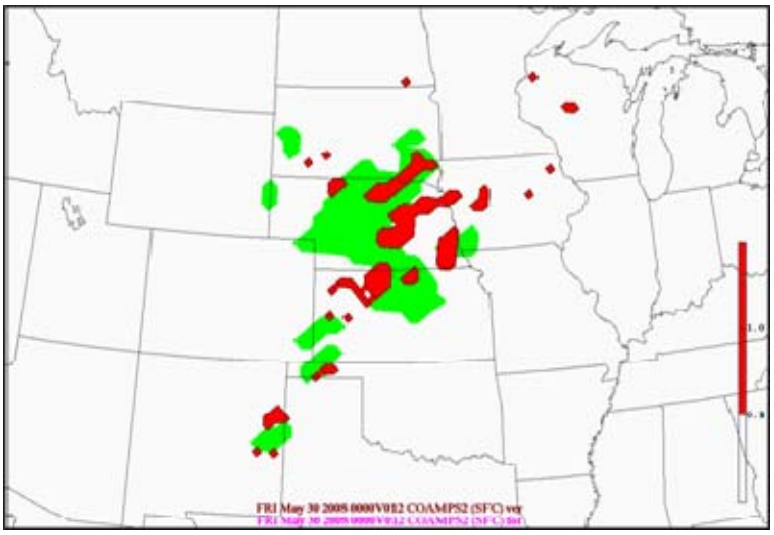


Figure 72. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter Theta-e and LR intersections (green area) and Verification (red area) (valid 29/2345Z – 30/0015Z).

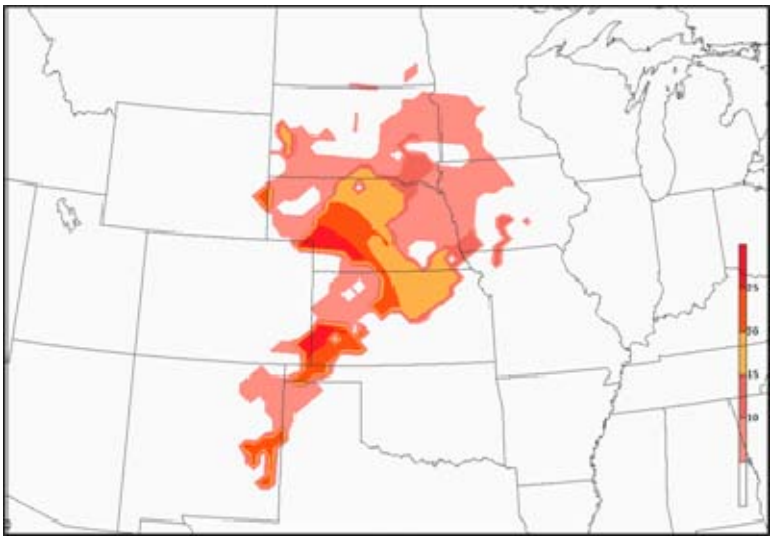


Figure 73. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter TELR Probabilistic Forecast (probabilities > 5% are shaded in 5% increments).

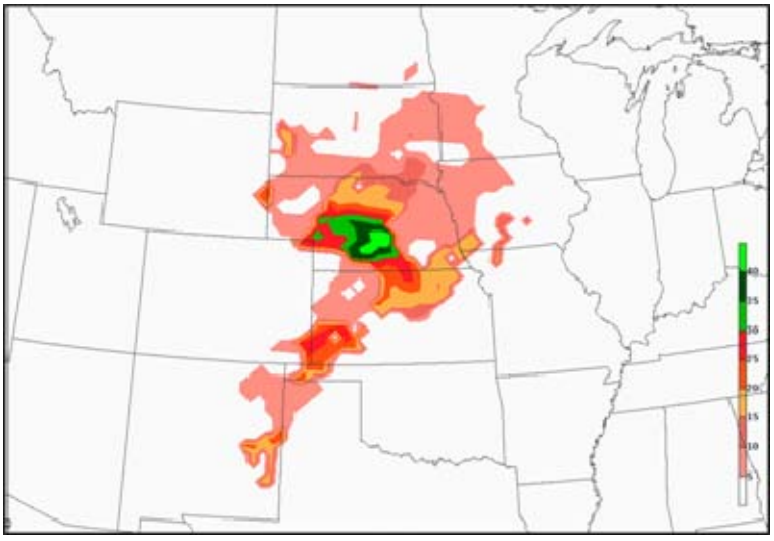


Figure 74. 080529NAM-WRF 12Z Run – 12-hr Forecast (valid 30/00Z). 10× Filter TELR^{+} Probabilistic Forecast (probabilities > 5% are shaded in 5% increments).

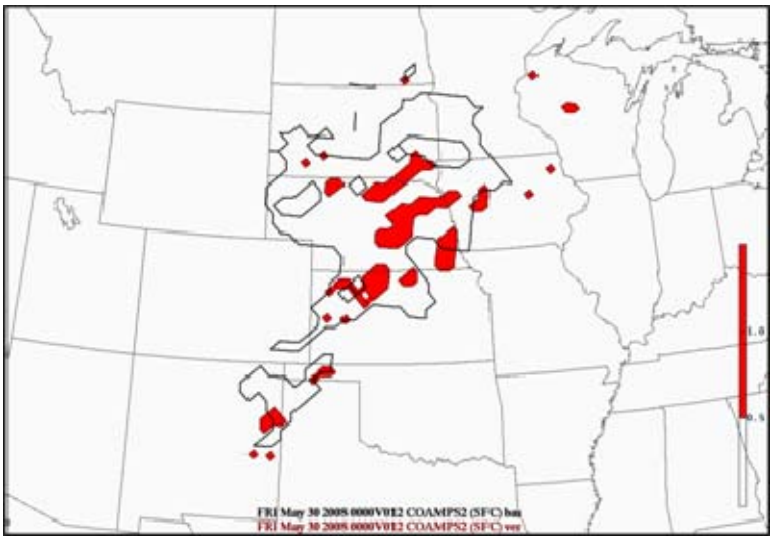


Figure 75. Verification (red area) using Radar Reflectivity Composites and Storm Reports and STR (black line) valid 080529/2345Z – 080530/0015Z.

Is there an increase in skill for the 10× filter forecast (or any other filter)? The BSS at each tau for the no filter, 10× filter, and 40× filter TELR and TELR^{+} forecasts are shown in Figure 76. Only the forecasts in the STR are considered in attempt to avoid skill inflation. There seems to be no significant difference in skill between the unfiltered and any of the other forecasts (will be expanded upon later in this chapter).

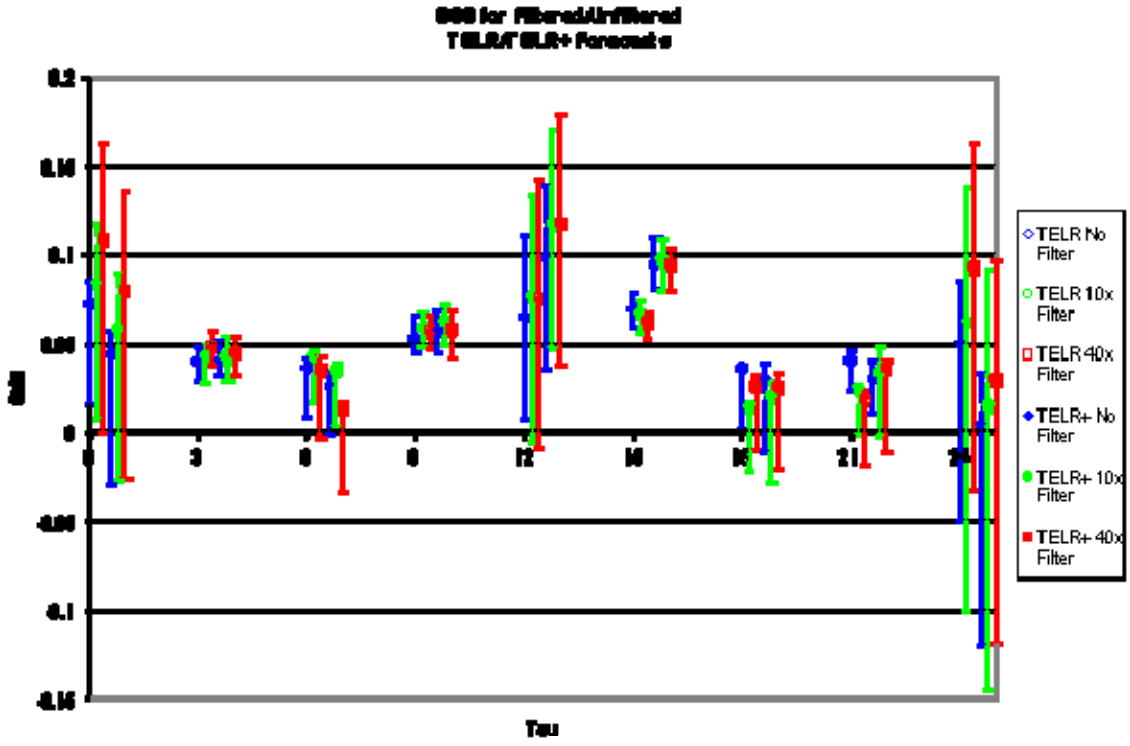
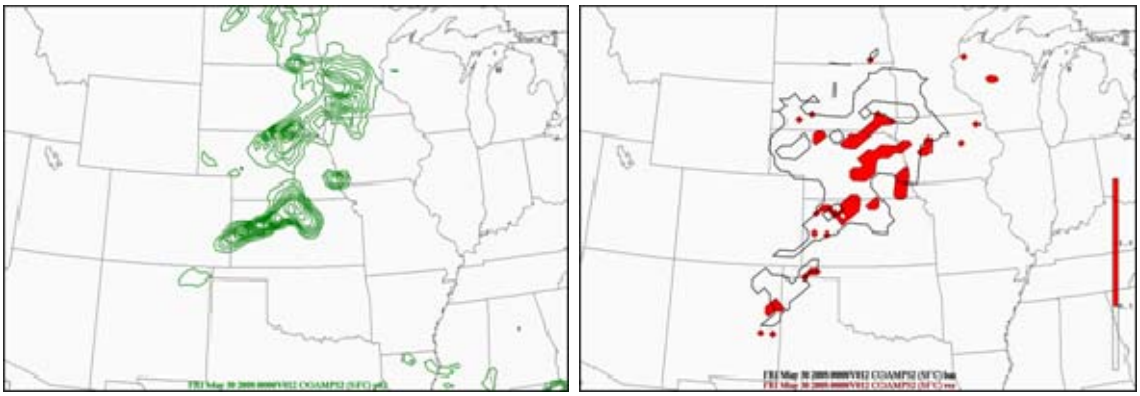


Figure 76. Brier Skill Score (BSS) for the 24-hr forecast period for the TELR and TELR^{+} Forecasts for only the Severe Threat Region (STR) gridpoints using No Filter (blue), 10×Filter (green) and 40×Filter (red) with 95% Confidence Intervals.

d. Other Potential Problems

One of the first potential problems of the severe forecast is the definition of the STR using ± 3-hr convective precipitation. As mentioned in Chapter III, the method produces an instantaneous probability of severe weather at the forecast tau, but the six-hour window will largely over estimate the region in which severe convection has potential (STR). This scenario is demonstrated in Figure 77 (a) and (b) at a forecast tau = 12. The convective precipitation valid between tau = 12 – 15 (+3-hr) is shown in Figure 77 (a) and the STR and verification is shown in Figure 77 (b).



(a) (b)

Figure 77. 080529NAM-WRF 12Z Run – (a) 15-hr Forecast (valid 30/00Z – 30/03Z) 3-hr Cumulative Convective Precipitation and (b) Verification (valid 30/00Z) (red area) and the STR (black line).

At tau = 12 (00Z), western Nebraska was under a meso-high as the convection was initiating southeastward along a surging outflow boundary (evident in Figure 78). The STR valid at 00Z, shown in Figure 77 (b), indicates western Nebraska as having a potential for severe weather, but this was not valid due to the meso-high. The +3-hr convective precipitation valid 30/0Z – 30/03Z, shown in Figure 77 (a), actually forecasted the general structure of the convection arching from northeastern Nebraska southwestward toward the Texas panhandle, shown in Figure 77 (b). This is one example of a possible error in the forecast procedure.

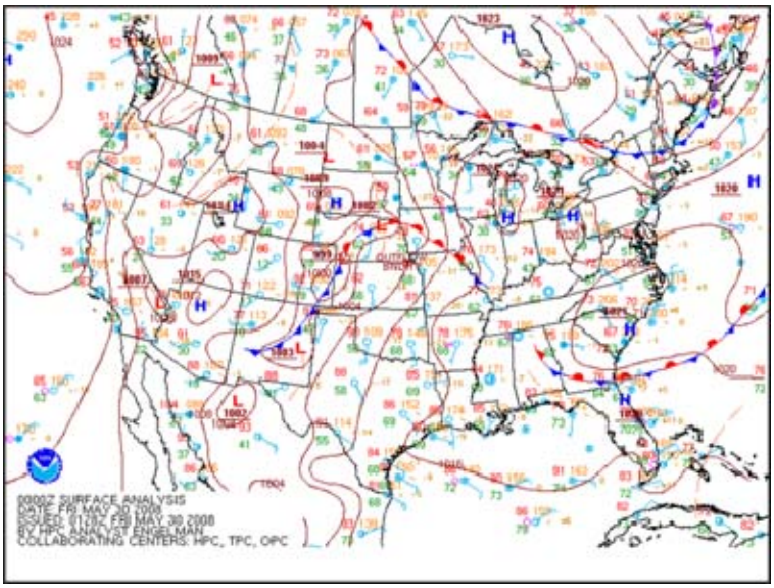


Figure 78. Surface Analysis valid at 30/00Z (From: SPC webpage, n.d.).

Another interesting finding from the running of multiple filters is the predictability issue. As larger wavelength features are used in the forecast process, the predictability should increase (Eckel, 2008) if there is any validity in the method. The model non-linear error growth should be faster for the shorter wavelength features, decreasing the predictability faster than for the longer wavelength features. The results showed no significant predictability difference between the filtered and non-filtered forecasts. There are a few possible explanations for this result.

First, both short and long wavelength feature forecasts show nearly positive significant skill throughout the forecast period. Possibly if the forecast were extended to 48 hrs, the crossover to no skill will more evident in the second day (24 – 48-hr) forecast. This does not seem to be the case though because both wavelength features forecasts have skill decreasing at nearly the same rate.

Another possible explanation for the nearly equal skill between the two wavelength forecasts is the verification procedure. Since the method verifies severe weather within a 96 × 96 km square (9216 km^2 or 3558mi^2), as model forecast error grows, the severe forecasts are still being verified because such a large verification area is

being used. In the following illustration (Figure 79), the red and blue circles indicate the short and long wavelength forecasts of severe weather and the red dot represents verification (the grid spacing is 32 km). Assuming both forecasts initialized well, the error growth of the mesoscale feature (red circle) has grown more than the synoptic scale feature (blue circle). The distance from the verification of the mesoscale feature is approximately 50 km, but the synoptic scale error is only 15 km.

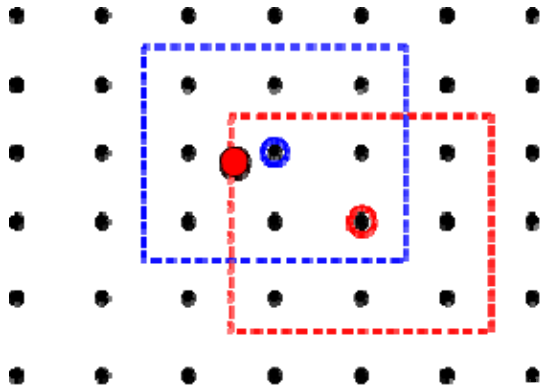


Figure 79. Example of Verification (red solid dot) using 96 × 96 km Box for Short (red circle and dotted box) and Long Wavelength Forecasts (blue circle and dotted box).

In this instance, both forecasts verified although the synoptic scale feature clearly had less error. Decreasing the verification region to a 32 × 32 km region centered on the gridpoint could potentially distinguish the predictability between the two scale forecasts.

Another reason that predictability is not shown to decrease with time could be the diurnal nature of convection. Since the method trained over all forecast hours and did not use forecast hour as a predictor, there is over and under prediction at each tau. This diurnal increase in convection in the afternoon results in more verified forecasts between 21 – 03Z, giving the method higher significant positive skill at tau = 09 and 15 than any other tau. The exception is the 12-hr forecast where the 95% confidence interval increased greatly. A possible fix to this problem would be to somehow remove the diurnal variation of skill to see the limit of predictability. Also, adding the forecast tau as a predictor would tend to remove the over and under forecasting.

THIS PAGE INTENTIONALLY LEFT BLANK

V. SUMMARY AND CONCLUSION

The use of theta-e ridges to forecast regions of severe weather is not a new concept, but LR ridges, on the other hand, are not used as frequently. The theory tested in this research states that intersections of these ridge axes maximizes the potential and conditional instabilities, respectively. To do this, the process was automated by mathematically defining a ridge axis and using a series of Boolean operators to determine if the ridge intersections were in a general severe threat region (STR). Linear discrimination analysis was performed using the operational NAM-WRF for dependent and independent datasets from the spring of 2008 severe season.

The statistical analysis using the four predictors (TELR) indicated significant positive skill at the analysis through the 21-hr forecast tau. Adding MUCAPE and 0 – 6 km shear to the predictors ($TELR^{+}$) appeared to introduce more of a diurnal correlation. One main feature noted was both forecasts (TELR and $TELR^{+}$) over forecasting of severe convection overnight into the morning hours and the under forecasting of severe during max heating. This is a clear indication that forecast hour should have been added as a predictor.

The forecast theta-e and LR ridge intersections inside the STR appeared to have mesoscale structure, but how true are these wavelength (160 km) features? Are these automated features the same a forecaster would indicate, or would other features be weighted differently. A point made in Chapter I was that we assumed the ridging of the theta-e and LR fields would give us insight to the advection characteristics. As the wavelength features become longer, the advections of theta-e correlate with the LLJ, but do the shorter wavelength features imply anything about advection? Will the larger scale features (meso-β to meso-α) have greater skill because of the intersections of these advective fields (i.e., LLJ and mid-level jet)? Also, the model error growth is larger for smaller scale features, so will the synoptic scale extend the predictability?

In an attempt to circumvent the potential error in the perturbations, filters were applied to the theta-e and LR fields to remove the ageostrophic structure. By filtering the

theta-e and LR fields, the meso-β to meso-α scale wavelength features dominated. The process was performed again and the statistical results were not significantly different than the unfiltered forecasts. By filtering the small-scale features, the predictability of the model forecast parameters should have increased due to the slow variability of the background.

A few reasons were addressed as to the potential reasons for the lack of skill and inability to determine a predictability limit such as removing diurnal trends and changing verification procedures.

1. Objectives Answered

(1) The forecast method using the four predictors (theta-e, LR, and the ridge magnitudes), the TELR forecast, has significant positive skill at the analysis hour indicating validity of method. The method appeared to have positive skill throughout the 21-hr forecast as well.

(2) The linear discrimination analysis was performed effectively as a forecast procedure training on the dependent dataset and making a probabilistic forecast using the independent dataset. There was a problem addressed in the delineation of the severe threat region (STR) by using the ± 3-hr convective precipitation for an instantaneous probabilistic forecast.

(3) The BSS, ROCSS, reliability diagrams and VS were used to determine skill of the method. The use of the reliability diagrams indicated over and under forecasting of severe weather at different forecast taus. This led to the conclusion of adding forecast tau as a predictor in the linear discrimination analysis for future research.

(4) The addition of MUCAPE and 0 – 6 km shear to the four previous predictors, the TELR$^+$ forecast, was an attempt to refine the forecast and potentially increase skill over the TELR forecast (four predictors). This forecast indicated lack of significant skill at the analysis (using 95% CI), but there appeared to be a larger diurnal correlation than with the TELR forecast. The same over and under forecasting problem appeared in the TELR$^+$ forecast as well.

(5) To understand if there was validity in the mesoscale features in the model forecast ridge axes, a number of filters were applied to the theta-e and LR fields to remove the perturbations and arrive at the background fields. The application of these filters created a forecast that appeared to be more typical of a hand analysis, but the skill was no different than the unfiltered forecast. This was addressed and there was a potential error in the verification procedure which masks any significant difference between the short and long wavelength forecasts.

In conclusion, this forecast could in fact be a useful tool for the forecasters at the 26th OWS to delineate regions of increased severe weather probability in the severe threat region (STR). A few issues need to be addressed to increase skill such as the addition of forecast tau to the predictors, the decrease in spatial coverage of the STR due to over forecast precipitation coverage, and a decrease in the verification coverage.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. FUTURE RESEARCH

A. DETERMINE VALIDITY OF METHOD USING ANALYSIS

In order to determine if a method has skill, it should be first tested on analysis and trained using analysis data. The hourly RUC analysis or 00Z analysis would serve as a reasonable representation of the atmosphere (reasonable is relative).

B. DIRECTIONAL SHEAR

Miller (1972) stated that veering between the low- and mid-levels is conducive to severe weather as well as the intersection of the mid- and lower-level jets. Weisman and Klemp (1984) also extended their idea of balancing shear/buoyancy (i.e., BRN) with clockwise curved wind hodographs. They learned from Wilhelmson and Klemp (1978) that in a unidirectionally sheared environment, storms split and form supercells that rotate cyclonically (anti-cyclonically) and turn to the right (left) of the mean flow. When Weisman and Klemp (1984) veered the vertical wind profile, as thunderstorms split, the new updraft growth was favored for the cyclonically rotating storm (right-mover) of the split pair.

Browning (1965) stated that "most severe local storms in the Midwest appear to travel anomalously to the right of the mean wind…and tornadoes are liable to occur." A detailed case study was performed on a right-turning supercell that produced large hail and funnel clouds in Oklahoma on 24 May 1968. The rawinsonde sampled storm wind environment revealed 180 degrees of cyclonic turning between 0 – 6 km (Barnes, 1970). This leads to a key question: Is there some way, qualitatively or quantitatively, to assess veering wind profiles by using the LR and theta-e ridges?

As stated before, the low-level theta-e ridge analysis represents the warm, moist air advective regime as shown in (d) in Figure 1. Let's assume the LR ridge analysis between 950 mb and 500 mb represents the mid- to low-level differential advection of cold air aloft/warm air beneath by their respective jets, shown in Figure 1 (a) and (b). We

can then assume the total LR ridge axis, shown in Figure 1 (c), represents the average low- to mid-level jet vector assuming jet speeds are comparable in magnitude. This is a somewhat safe assumption with the typical mean low- and mid-level jet speeds for central US baroclinic systems approximately 40 and 50 kts respectively (Bonner, 1968; Miller, 1972). The intersection angle between these ridges, shown Figure 1 (e), should give us information of the veering wind profile assuming a southerly component to the low-level wind and a westerly component to the mid-level wind. As Miller (1972) stated, 30 degrees of cyclonic turning of the horizontal winds with height increases the certainty of severe development. Miller (1972) also stated that the intersection of the low- and mid-level jets is a preferred zone for convective development. The intersection of the LR and theta-e ridges is also an approximation for the location of the intersection of the low- and mid-level jet axis as shown in Figure 1 (e).

Does the angle of intersection between the theta-e and LR ridge axes give insight to the storm mode of types of severe weather to be produced? Hand analysis of each intersection can be cumbersome, but maybe each ridge axis can be assigned a slope in the x-y direction. This would allow the slope of each ridge intersection to be compared and determine the angle of intersection. This angle of intersection could be an addition to the predictors and potentially discriminate between severe weather types using multiple linear discrimination analysis (Wilks, 1995).

APPENDIX. LIST OF FORMULAS

(1) Relative Operating Characteristic Skill Score (ROCSS):

$$ROCSS = \frac{A - A_{rand}}{A_{perf} - A_{rand}} = \frac{A - 1/2}{1 - 1/2} = 2A - 1$$

$A =$ the area under the curve traced by the ROC diagram (shown below as the hatched region in Figure 80).

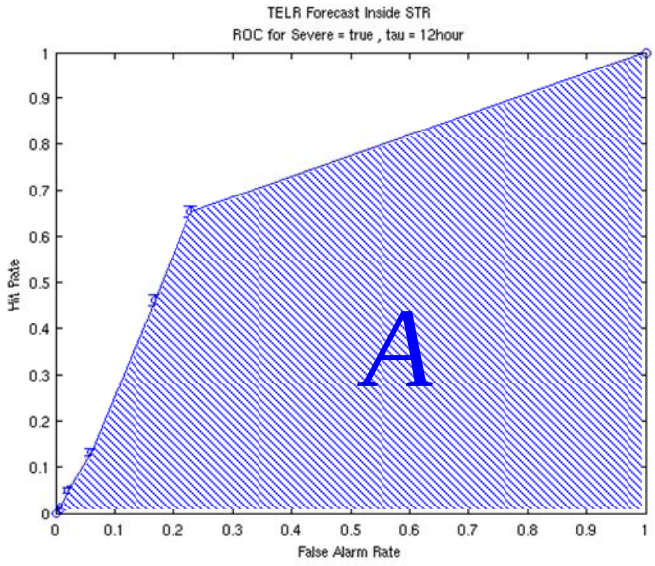


Figure 80. ROC Diagram (tau = 12) for the TELR Forecast Inside the Severe Threat Region (STR) (hatched region indicates A).

(2) Brier Skill Score (BSS):

$$BSS = \frac{BS - BS_{ref}}{0 - BS_{ref}} = 1 - \frac{BS}{BS_{ref}}$$

Where,

$$BS = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2 ,$$

$$BS_{ref} = BS_{\bar{o}} = \frac{1}{n}\sum_{k=1}^{n}(\bar{o} - o_k)^2 ,$$

$$o_k = \text{probabilistic forecast},$$

$$\bar{o} = \text{sample climatology} = \frac{1}{n}\sum_{k=1}^{n} o_k ,$$

$$\text{and } n = \text{number of forecasts}.$$

(3) BSS calculated by its decomposed components (reliability, resolution and uncertainty) derived from the reliability diagram:

$$BS = \text{"Reliability" term} + \text{"Resolution" term} + \text{"Uncertainty" term},$$

$$\text{"Reliability" term} = \frac{1}{n}\sum_{i=1}^{I} N_i (y_i - \bar{o}_i)^2 ,$$

$$\text{"Resolution" term} = \frac{1}{n}\sum_{i=1}^{I} N_i (\bar{o}_i - \bar{o})^2 ,$$

$$\text{"Uncertainty" term} = \bar{o}(1 - \bar{o}),$$

$$I = \text{number of forecast taus}.$$

(4) Value Score (VS):

$$VS = \frac{EE_{\mathrm{f}} - EE_{\mathrm{clim}}}{EE_{\mathrm{perf}} - EE_{\mathrm{clim}}},$$

$$EE_{\mathrm{f}} = (\mathrm{p}_{1,1} + \mathrm{p}_{1,0})C + \mathrm{p}_{0,1}L = C\sum_{j=0}^{1}\sum \mathrm{p}(y_i, o_j) + L\sum_{i<D} \mathrm{p}(y_i, o_1),$$

$$EE_{\mathrm{clim}} = \begin{cases} C, & \text{if C/L} < \bar{o} \\ \bar{o}L, & \text{otherwise} \end{cases},$$

$$EE_{\mathrm{perf}} = \bar{o}C,$$

The variables $p_{1,1}$, $p_{1,0}$ are as follows: $p_{1,1}$ = the joint frequency that the forecast y_i is above the decision threshold and the event occurs, and $p_{1,0}$ = the joint frequency that the forecast is above the probability threshold but the event does not occur.

C = cost = cost to take protective action in preparation for adverse event.

L = loss = cost if no protective action is taken and the adverse event occurs.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Atkins, N. T., & Wakimoto, R. M. (1991). Wet microburst activity over the southeastern United States: Implications for forecasting. *Wea. Forecasting*, 6, 470-482.

Barnes, S. L. (1970). Some aspects of severe, right-moving thunderstorms deduced from mesonet rawindsone observations. *J. Atmos. Sci.*, 27, 634-648.

Brooks, H. E., & Doswell, C. A. III. (1993). Extreme winds in high precipitation supercells. Preprints, *17th Conf. on Severe Local Storms*. St. Louis, MO, Amer. Meteor. Soc., 173-177.

Brooks, H. E., Doswell C. A. III, & Cooper, J. (1994). On the environments of tornadic and nontornadic mesocyclones. *Wea. Forecasting*, 9, 606-618.

Browning, K. A. (1964). Airflow and precipitation trajectories within severe local storms which travel to the right of the winds. *J. Atmos. Sci.*, 21, 634-639.

———. (1965). The evolution of tornadic storms. *J. Atmos. Sci.*, 22, 664-668.

Bonner, W. D. (1968). Climatology of the low-level jet. *Mon. Wea. Rev.*, 96, 833-850.

Chen, C.-H., & Orville, H. D. (1979). Effects of mesoscale convergence on cloud convection. *J. Appl. Meteor.*, 19, 256-274.

COMET Program. (2006). *Operational Models Matrix: Characteristics of NWP and Related Forecast Models – Convective Parameterization Schemes – Betts-Miller-Janjic.'* Retrieved March 18, 2009, from http://www.meted.ucar.edu/nwp/pcu2/mancp1.htm

Davies, J. M., and Johns, R. H. (1993). Some wind and instability parameters associated with strong and violent tornadoes. *The Tornado: Its Structure, Dynamics, Prediction, and Hazards, Geophys. Monogr.* No. 79, Amer. Geophys. Union, 573-582.

Doswell, C. A. (1987). The distinction between large-scale and mesoscale contribution to severe convection: A case study example. *Wea. Forecasting*, 2, 3-16.

———. (2001). Severe Convective Storms – An Overview. *Severe Convective Storms, Meteor. Monogr.*, No. 50, Amer. Meteor. Soc., 1-26.

Doswell, C. A., & Bosard, L. F. (2001). Extratropical synoptic-scale processes and severe convection. *Severe Convective Storms, Meteor. Monogr.* No. 50, Amer. Meteor. Soc., 27-69.

Eckel, F. A., (2008). *Introduction to Ensemble Forecasting*. Naval Postgraduate School Lecture Notes for MR4324 and MR4325.

Fawbush, E. J., & Miller, R. C. (1954). The types of air masses in which North American tornadoes form. *Bull. Amer. Meteor. Soc.*, 35, 154-165.

Foster, D. F., & Bates, F. (1956). A hail size forecasting technique. *Bull. Amer. Meteor. Soc.*, 37, 135-140.

Heideman, K. F., Stewart, T. R., Moninger, W. R., & Reagan-Cirincione, P. (1992). The weather information and skill experiment (WISE): The effect of varying levels of information on forecasting skill. *Wea. Forecasting*, 8, 25-36.

Igau, R. C., & Nielson-Gammon, J. W. (1998). Low-level jet development during a numerically simulated return flow event. *Mon. Wea. Rev.*, 126, 2972-2990.

Janjic, Z. I. (1994). The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev*., 122, 93 – 106.

Jirak, I. L., & Cotton, W. R. (2006). Observational analysis of the predictability of mesoscale convective systems. *Wea. Forecasting*, 22, 813-838.

Johns, R. H., Davies, J. M., & Leftwich P. W. (1993). Some wind and instability parameters associated with strong and violent tornadoes, 2. Variations in the combinations of wind and instability parameters. *The Tornado: Its Structure, Dynamics, Prediction and Hazards, Geophys. Monogr.,* No. 79, Amer. Geophys. Union, 611-624.

Lenning, E., Fuelberg H. E., & Watson A. I., (1998). An evaluation of WSR-88D severe hail algorithms along the northeastern gulf coast. *Wea. Forecasting*, 13, 1029 – 1045.

Maddox, R. A. (1980). Mesoscale convective complexes. *Bull. Amer. Meteor. Soc.*, 61, 1374-1387.

McNulty, R. P. (1995). Severe and convective weather: A central region forecasting challenge. *Wea. Forecasting*, 10, 187-202.

McNulty, R. P. (1981). A statistical approach to short-term thunderstorm outlooks. *J. App. Meteor.*, 20, 765 – 771.

Miller, R. C. (1962). Statistical prediction by discriminate analysis. *Met. Monogr.*, 4, No. 25. Amer. Meteor. Soc.

Miller, R. C. (1972). Notes on analysis and severe-storm forecasting procedures of the Air Force Global Weather Central. AWS Tech. Rep. 200 (rev). Air Weather Service, Scott AFB, IL.

NCEP website: Accessed March 18, 2009, from http://www.emc.ncep.noaa.gov/mmb/papers/manikin/2/prdgen_documention.txt

Pagliaro, D. E. (2008). Verification of the AFWA 3-element severe weather forecasting system. M.S. thesis, Naval Postgraduate School, Monterey, CA.

Ray, P.S., Ed. (1986). *Mesoscale Meteorology and Forecasting*. Amer. Meteor. Soc.

Rossby, C.-G. (1932). Thermodynamics applied to air analysis. *MIT Meteorological Papers*, Vol. 1, No. 3, 31-48, plus 11 plates.

Rotunno, R. (1993). Supercell thunderstorm modeling and theory. *The Tornado: Its Structure, Dynamics, Prediction, and Hazards, Geophys. Monogr.*, No. 79, Amer. Geophys. Union, 57-73.

Schultz, D. M., Schumacher, P. N., & Doswell III, C. A.. (2000). The intricacies of instabilities. *Mon. Wea. Rev*., 128, 4143-4148.

Sherwood, S. C. (2000). On moist instability. *Mon. Wea. Rev.*, 128, 4139-4142.

SPC website: Accessed March 18, 2009 from http://www.spc.noaa.gov/

Thompson, R. L., Lewis, J. M. & Maddox, R. A. (1994). Autumnal return of tropical air to the Gulf of Mexico's coastal plain. *Wea. Forecasting*, 9, 348-360.

Weisman, M. L., & Klemp, J. B. (1982). The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, 110, 504-520.

———. (1984). The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, 112, 2479-2498.

———. (1986). Characteristics of isolated convective storms. *Mesoscale Meteorology and Forecasting*, P. S. Ray, Ed., Amer. Meteor. Soc.. 331-358.

Wilhelmson, R. B. & Klemp, J. B. (1978). A numerical study of storm splitting that leads to long lived storms. *J. Atmos. Sci.*, 35, 1974-1986.

Wilks. D. S. (1995). *Statistical Methods in the Atmospheric Sciences.* Academic Press, p. 467.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. 26th Operational Weather Squadron
   Barksdale AFB, Louisiana

4. Storm Prediction Center
   Norman, Oklahoma

5. Air Force Weather Technical Library
   Asheville, North Carolina